NBS SPECIAL PUBLICATION 343

Precision Measurement and Fundamental Constants

U.S. DEPARTMENT OF COMMERCE
National Bureau of Standards

UNITED STATES DEPARTMENT OF COMMERCE · Maurice H. Stans, *Secretary*
NATIONAL BUREAU OF STANDARDS · Lewis M. Branscomb, *Director*

Precision Measurement and Fundamental Constants

Proceedings of the International Conference held at the National Bureau of Standards Gaithersburg, Maryland, August 3–7, 1970

Edited by

D. N. Langenberg

Department of Physics
University of Pennsylvania
Philadelphia, Pennsylvania 19104

and

B. N. Taylor

Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

National Bureau of Standards Special Publication 343

Nat. Bur. Stand. (U.S.), Spec. Publ. 343, 543 pages (August 1971)
CODEN: XNBSA

Issued August 1971

Abstract

This volume presents the Proceedings of the International Conference on Precision Measurement and Fundamental Constants, held at the National Bureau of Standards in Gaithersburg, Maryland, from August 3 through August 7, 1970. The conference brought together theoretical, experimental, and applied scientists for the purpose of discussing modern techniques of precision physical measurement and their application, along with modern theoretical developments, to the determination of the fundamental constants. The topics covered were: frequency and time standards; length standards; the velocity of light; the Rydberg constant; electrical standards; the proton gyromagnetic ratio; the Faraday constant; atomic masses; the proton magnetic moment; Josephson effects; x-rays; fine and hyperfine structure in simple atoms; lepton *g*-factor anomalies; the gravitational constants; least squares adjustments of the constants. These Proceedings contain the invited tutorial talks as well as the contributed papers presented at the Conference. Also included are the post-paper discussions and a panel discussion entitled "Should Least Squares Adjustments of the Fundamental Constants Be Abolished?"

Key words: Fundamental constants; least squares adjustments; precision measurements.

Library of Congress Catalog Card Number: 73-610425

Foreword

This Conference brought together, on an international basis, scientists actively engaged in experimental and theoretical research on precision measurements relating to the fundamental physical constants. We believe that this is the first time that this has been done in such a deliberate and comprehensive way.

The National Bureau of Standards was especially pleased to act as host to this Conference because of our role, which we share with other national standards laboratories, for the development, maintenance and dissemination of the units of physical measurement. The essential connection between the fundamental physical constants and our choice and definition of the base units in our scheme of physical measurement has been apparent for many years. Moreover, it is scientific invention and discovery which provide us new opportunities for strengthening these connections; examples from recent years include coherent light sources and macroscopic quantum phase coherence in superconductors. This is not a new idea. Indeed, in a larger sense it has been the goal for over 150 years of men devoted to the science of measurement to develop a "philosophically true" system "established on an invariant base, fixed in nature" to which we could turn "repeatedly to find the means of verification."[1]

I should again emphasize that the main objective of the Conference was to discuss modern techniques of precision physical measurement and their application, along with modern theoretical developments, to the determination of the fundamental constants. It was not the purpose of the Conference to arrive at new adjusted values of the constants. We sincerely hope, however, that these Proceedings will provide both a valuable source of information for those making future adjustments, and a stimulus for further research.

The National Bureau of Standards is pleased to have been able to join with the International Union of Pure and Applied Physics, the International Bureau of Weights and Measures, the Committee on Data for Science and Technology of the International Council of Scientific Unions, and the Committee on Fundamental Constants of the U.S. National Academy of Sciences-National Academy of Engineering-National Research Council, in co-sponsoring this Conference.

E. Ambler, Director

Institute for Basic Standards
National Bureau of Standards

[1] J. H. Van Swinden "Précis des Operations; Système Métrique" le 15 messidor an 7 (French Revolutionary Calendar), July 3, 1799 (Gregorian Calendar).

Preface

Precision measurement plays an essential and crucial role in the development of all science and technology. Scientific theories cannot be tested, instruments and machines cannot be designed and built, and even the daily routine of the market place cannot proceed without precise quantitative measurement. The demand for precision and accuracy is greatest at the frontiers of science and technology. Here our increasing understanding of the fundamental nature of the universe and our increasing ability to put our knowledge to practical use depends on continuing refinement of our theories and of the technology used in testing these theories against reality.

In all of our fundamental theories there appear a certain few parameters which characterize the fundamental particles and interactions we find in nature. These are the fundamental physical constants. The precision determination of the numerical values of these constants has long been and remains one of the principal objectives of experimental science. This is not because there is any intrinsic virtue in accumulating lists of ever longer numbers, but because the fundamental constants are the quantitative links between our most basic theories and the physical reality we wish them to describe. Our theories must stand or fall according to their ability to make quantitative predictions which agree with experimental observations to the maximum accuracy possible. The progress of our understanding of the physical world is therefore very much intertwined with the advance of the art of precision measurement and its application to the determination of the fundamental physical constants.

In recent years there have been notable advances in both theory and experiment relevant to the fundamental constants. New technologies have made possible new levels of precision: Computers have made possible more complex calculations. The laser and the maser have found application in many different experiments. The discovery of the Josephson effects in superconductors and superfluid helium has provided new ways to determine several important fundamental constants. There has been steady progress in precision measurements of the properties of fundamental particles and simple atomic systems. Current theoretical and experimental efforts in this field span almost the whole of physics, from elementary particle theory to experimental solid state physics. The International Conference on Precision Measurement and Fundamental Constants, held at the National Bureau of Standards, Gaithersburg, Md., August 3–7, 1970, was organized to bring together active workers from all these diverse fields to discuss these recent developments, to compare notes on the present state of the art in their specialized areas, and to consider what may be the future course of development in the field. This volume contains the invited and contributed papers given at the Conference, together with the accompanying discussion.

The Conference sessions were organized around particular fundamental constants or types of experiment or theory. This arrangement has been continued in these Proceedings. They begin with a transcription of the opening remarks made by J. Terrien, Director of the Bureau International des Poids et Mesures, and a keynote address by L. M. Branscomb, Director of the National Bureau of Standards. There are twenty invited review papers. These were intended to convey a picture of the past history, present status, and likely future development of each area. Accompanying these are contributed papers describing current work by active workers in the field. Taken together, the invited and contributed papers and the accompanying discussion provide a rather complete picture of the field as it stood in mid-1970.

The critical analysis and assessment of all available experimental and theoretical information and the distillation from these data of a complete and consistent set of fundamental constants is an essential and extremely important aspect of the field of precision physical measurements. It is a subject, like politics and the weather, about which everyone seems to have strong opinions, whatever

his specialty. It was recognized by the organizers at the outset that any attempt to achieve a Conference consensus on the best methods of assessing and handling data or the likely best values of the constants based on data presented at the Conference was totally out of the question. (The editors' own experience indicates that it is nearly impossible to get such a consensus among three people, much less several hundred.) However, the organizers felt it would be probably fruitful and certainly entertaining to provide a forum for the expression of personal views on the subject of fundamental constants adjustments. A panel discussion was therefore planned under the title, "Should Least Squares Adjustments of the Fundamental Constants be Abolished?" It featured prepared contributions from six active partisans and led to a lively discussion among the Conference participants. The discussion is reproduced *in toto* here.

One unique feature of the Conference was the emergence of the limerick as a popular mode of expression. By the end of the Conference we had collected a number of examples of this literary art form. The reader will encounter them here and there in this volume. Unlike the conventional scientific prose, they have not been reviewed for scientific and literary quality.

It is sometimes taken as a truism in the field of precision measurement and fundamental constants that each generation of physicists advances the general level of precision and accuracy by about one order of magnitude. Insofar as it is possible at all to characterize the attitudes of as diverse a group of conferees as attended this Conference, it may be said that the pace seems to be accelerating. It has not been long since uncertainties of several tens of parts per million were the rule in most of the best precision measurements. Now workers compelled to quote uncertainties larger than several parts per million sometimes seem apologetic, and sub-part-per-million uncertainties are becoming common. The full exploitation of the many new techniques discussed at this Conference should take our knowledge of the fundamental constants and our precision measurement technology well beyond the part per million level. But one wonders whether this is really as far as it will be possible to go. Will as yet undreamed of advances ultimately give us, say, the fine structure constant to one part in 10^{10}, together with the theoretical understanding necessary to make such an accuracy useful? Short of a transmission from some extraterrestrial higher civilization, the answer to this question can only be supplied by continuing applications of ingenuity and persistence. We hope we will be able to join most of the Conference attendees in finding in the Proceedings of a similar conference at the turn of the century a suitable postscript to the following limerick:

Take an "e" and an "h" and a "c",
Combine them dimensionlessly.
Then Eddington's number
From one over alpha
Leaves zero, three six and query.

G. W. Series

We would like to express our thanks to the members of the Program and Editorial Committees for their generous contributions to the production of these Proceedings. We should also like to thank the many organizations and individuals who assisted in making the Conference the successful venture it was. These include the sponsors, contributors, and committee members listed on the following pages, together with numerous members of the National Bureau of Standards staff, most especially R. T. Cook and J. A. Suddeth.

D. N. Langenberg

B. N. Taylor

CONFERENCE COMMITTEES

General Chairman

L. M. Branscomb

Executive Committee

E. Ambler (Chairman)
E. R. Cohen
K. G. Kessler
D. N. Langenberg
B. N. Taylor
J. S. Thomsen

Program Committee

P. L. Bender
M. Danos
R. D. Deslattes
V. W. Hughes
D. N. Langenberg (Chairman)
K. R. Lea
D. R. Lide
B. N. Taylor
J. S. Thomsen

International Advisory Committee

K. M. Baird, Canada
J. A. Bearden, U.S.A.
A. H. Cook, U.K.
H. R. Crane, U.S.A.
R. H. Dicke, U.S.A.
S. D. Drell, U.S.A.
H. E. Duckworth, Canada
J. M. Dunworth, U.K.
L. Essen, U.K.
W. M. Fairbank, U.S.A.
P. Franken, U.S.A.
G. Herzberg, Canada
V. W. Hughes, U.S.A.
A. Javan, U.S.A.
M. Kotani, Japan
W. E. Lamb, Jr., U.S.A.
F. J. Lehany, Australia
J. E. Mercereau, U.S.A.
R. Novick, U.S.A.
H. Preston-Thomas, Canada
N. F. Ramsey, U.S.A.
F. D. Rossini, U.S.A.
G. W. Series, U.K.
U. Stille, West Germany
J. Terrien, France
C. H. Townes, U.S.A.
A. H. Wapstra, Netherlands
T. Yamamoto, Japan

Editorial Committee

J. A. Barnes
P. L. Bender
R. D. Cutkosky
M. Danos
R. D. Deslattes
V. H. Dibeler
W. J. Hamer
F. K. Harris
D. N. Langenberg (Coeditor)
K. R. Lea
D. R. Lide
W. C. Martin
C. H. Petersen
B. W. Steiner
D. R. Tate
B. N. Taylor (Coeditor)
J. S. Thomsen

Sponsors

International Union of
Pure and Applied Physics

Bureau International des
Poids et Mesures

Committee on Data for Science and
Technology of the International
Council of Scientific Unions

Committee on Fundamental Constants of the
U.S. National Academy of Sciences–
National Academy of Engineering–
National Research Council

National Bureau of Standards
U.S. Department of Commerce

Contributors

The International Conference on Precision Measurement and Fundamental Constants expresses its deep appreciation to the following organizations for their generous financial assistance which contributed greatly to making this meeting possible.

General Motors Corporation
Hewlett-Packard Company
International Business Machines Corporation
International Union of Pure and Applied Physics
Leeds & Northrup Company
National Conference of Standards Laboratories
National Science Foundation
North American Rockwell Corporation
Princeton Applied Research Corporation
RCA Corporation, David Sarnoff Research Center
Scientific Apparatus Makers Association
Texas Instruments Incorporated

Some of the participants at the International Conference on Precision Measurement and Fundamental Constants, held August 3–7, 1970, at the National Bureau of Standards, Gaithersburg, Maryland.

CONTENTS

PROTON GYROMAGNETIC RATIO

FARADAY CONSTANT AND ATOMIC MASSES

PROTON MAGNETIC MOMENT

SPECIAL GENERAL INTEREST TALK

JOSEPHSON EFFECT

X-RAYS

FINE AND HYPERFINE STRUCTURE IN SIMPLE ATOMS

LEPTON g-FACTOR ANOMALIES

GRAVITATIONAL CONSTANTS

PANEL DISCUSSION: SHOULD LEAST SQUARES ADJUSTMENTS OF THE FUNDAMENTAL CONSTANTS BE ABOLISHED?

Opening Remarks

J. Terrien

Directeur, Bureau International des Poids et Mesures, Sèvres, France

Mesdames, Messieurs:

C'est un grand honneur pour moi de présider cétte séance d'ouverture. Je profite de cette occasion pour remercier tres sincèrement le National Bureau of Standards de l'aide et de la coopération qu'il apporte au Bureau International des Poids et Mesures. Je désire aussi féliciter le National Bureau of Standards pour donner accueil à cette conférence qui est vraiment venue au moment le plus opportun.

I think I do not have to apologize for speaking of it in French because this conference is truly an international one. I think also that the subject matters which will be dealt with are quite timely. Not very many years ago, precision measurements at the higher level were measurements between standards which were of the same kind, for instance, comparing metre bar values and calibrating length standards. But now in a few years there is some difference. The best measurement standards now come from atomic physics so that they are more and more related to fundamental constants. As a consequence, the bringing together of precision measurements and fundamental constants is really an attractive matter to deal with and I am sure you will enjoy this Conference. And also, may I remark that perhaps some ill informed people in the past could believe that metrology was something of a lower level, not as attractive as pure physics. But I am sure that if you are not yet convinced of it, you will be convinced, after attending this Conference, that metrologists are not impure physicists.

KEYNOTE ADDRESS

Measurement Standards, Language of Discovery

L. M. Branscomb

Director, National Bureau of Standards

This International Conference is unique in many respects. To begin with it is held in the uniquely least pleasant season of the year. Washington is not known as a summer resort, and I can only hope that in addition to this laboratory, your bus and hotel are air conditioned. But not everyone finds the Washington climate so bad. Congressman Davis was here the other day and said that "Washington summers are not so bad, once you get used to the stupidity."

This conference may also be the occasion for a first visit to our new Gaithersburg laboratories for many of you. I hope you will take the time to explore the laboratories.

With regard to subject matter, this conference does follow a number of earlier ones that to some extent covered similar ground. There have been meetings sponsored by the IUPAP Commission on Atomic Masses and Related Constants. More recently the International Atomic Physics Conferences of 1968 in New York [1] and 1970 in Oxford dealt with many of the physics problems related to the fundamental constants, as did the Sommerfeld Centennial Conference on the Physics of One and Two Electron Atoms [2] in Munich in 1968. Perhaps the most similar meeting was the Avagadro Centennial Conference on the Physical Constants held in Turin in 1957, with proceedings published in Nuovo Cimento [3].

But none of these earlier conferences brought together the three groups we have here this week: the theorists, exploring basic processes in the framework of quantum electrodynamics and the nature of the superfluid state; the experimentalists interested in precision measurement of fundamental properties; and the physicists concerned with perfecting the scientific measurement system. All three groups share an active concern for the improvement of our knowledge of the fundamental constants of nature—and the continual probing of the limits to which we know them to be invariant in time and space. Without the help and advice of all three, an analysis of the most probable values of the constants is not possible.

Even more important, no one of these three groups can succeed without the other two for reasons that are inherent in the way in which physics makes progress. Superficially, of course, quantum electrodynamicists require exceedingly precise and accurate measurements of properties calculable in terms of the constants, such as Lamb shifts, lepton magnetic moments, hyperfine structure and so forth, in order to verify their higher order correction calculations and to test their models for proton structure and the like. Or, if you prefer to take the other point of view, a physicist wishing to deduce the best values of the constants from precision experiments cannot do so with Maxwell's equations and classical mechanics alone; he needs the now well-verified theory of quantum electrodynamics and can call up as many higher order corrections as he likes and can find money to pay for his friendly theorist's computer time and summer salary—for several summers I suspect. Of course the theorist can pursue his investigations using the fundamental constants as parameters of his theory—assuming them to be perfectly invariant—without ever knowing how anybody would realize in the laboratory the international absolute ohm. But, he still cannot settle disputes with his colleagues who prefer different models unless there is an experiment with which to compare, and which has the needed accuracy and reliability to make the comparison meaningful.

This raises right away the interesting question of what makes a comparison between experiment and theory meaningful, and I'll come back to that in a moment. But first let me note how intimately the fundamental UNITS and the fundamental CONSTANTS are interrelated. If we start out without any knowledge of physics and have no idea what might be a set of natural invariant parameters in terms of which we can describe physical phenomena, we have to start by picking a few gages that seem reasonably invariant to human experience. The period of a pendulum, and length and mass of two platinum-iridium objects, and the temperature interval between two phase transitions in water seemed like a good place to start. With these tools to create a measurement language capable of independent empirical reproduction one laboratory can share experience with another using only a written language to do so. And a large variety of physical experience can be described in terms of a measurement language with a large vocabulary based on a small number of root words each with a reasonably unambiguous operational meaning.

One purpose of this conference, then, is to discuss

how well we do define, realize and transfer the basic units of our measurement language—the language in terms of which a meaningful discussion of the constants—and therefore the basic questions of physics—must be conducted. This is not a matter to be left to the metrologists, as though an administrative detail and not a matter of physics. The following expressions show how the reciprocal fine structure constant is realized in operational terms in the recent Taylor, Parker, and Langenberg review [4]:

$$\alpha^{-1}=\left[\frac{1}{4R_\infty}\cdot\frac{c\Omega_{\mathrm{ABS}}}{\Omega_{\mathrm{NBS}}}\cdot\frac{\mu_p'}{\mu_B}\cdot\frac{2e/h}{\gamma_p'}\right]^{1/2}$$

$$\alpha^{-1}=\left[\frac{M_p^*}{4R_\infty}\cdot\frac{c\Omega_{\mathrm{ABS}}}{\Omega_{\mathrm{NBS}}}\cdot\frac{\mu_p'}{\mu_B}\cdot\frac{2e/h}{K^2F(\mu_p'/\mu_n)}\right]^{1/2}$$

$$K=A_{\mathrm{NBS}}/A_{\mathrm{ABS}}$$

These expressions could have been written with all the quantities observed in the laboratory specified in internationally defined base units, and the problem of realizing the ohm or the ampere would not appear explicitly. But, in fact, no investigator in this country can make such a measurement. To the extent that accurate values of the fine structure constant are needed for the further development of electrodynamics, for the exploration of the possible influence of strong forces on lepton properties or for explorations of the proton charge distribution, the accuracy with which we can realize and transfer the ohm or the ampere is just as important as the accuracy with which we measure the Rydberg or the ratio $2e/h$ from the ae Josephson effect. Physicists interested in a good value for alpha should be equally interested in all the physics that bears on the answer.

Yet this is not always easy. In 1963 a two day working session at BIPM attempting an evaluation of our knowledge of the gyromagnetic ratio of the proton spent a day and a half trying to track down and understand how the ampere in which one measurement was made related to the international ampere and thus to experiments in other countries. If we permit ourselves to be sloppy about our measurement language, especially if we do so internationally, progress in the exploration of physics will be thwarted.

I am reminded of a recent meeting of the International Organization for Standardization at which the U.S. and U.K. delegations almost came to blows over a proposed parliamentary action to table a motion. The British were in favor and the Americans against. Just before the delegates reached the shirt-sleeve-rolling-up stage, someone realized that in the United States to table a motion is to kill it, while in Britain the motion is brought to the table for action.

The accuracy with which basic units must be realized is, then, a function of demands placed upon them by the tests of basic physics made in terms of them. Since the progress of knowledge in basic physics is dependent upon the availability of this accuracy, the two are locked together in mutual interdependence. There is actually a three-way interdependence between the capabilities of the measurement system, the knowledge of the invariant properties of the physical world and the march of measurement technology. Using the measurement system, physics makes progress both with understanding and with new technology. Using the understanding and the technology we improve the measurement system (for example by the adoption of atomic standards for length and time interval). And simultaneously, applied science takes the technology and puts it to use—for example by using atomic clocks for pulse code modulation communications or for ship and aircraft navigation.

At NBS we are sometimes challenged to defend our plans for improvements of the measurement system in terms of the identified requirements of technology. And indeed, sometimes a new idea for improving the apex of the measurement system, such as the possible adoption of a stabilized laser for a realization of the meter, is picked up and translated in technology before the idea is fully perfected. But since the measurement system must be more accurately realized than the accuracy requirement for technological measurement compatibility, to base measurement research on known technological applications is to insure an inadequate and out of date measurement system.

We feel we have no choice but to base our efforts to improve the measurement system on a combination of scientific needs and scientific opportunity. (Scientific opportunity is a bureaucratic euphemism for having a good idea nobody else has had yet.) Thus, we at NBS are keenly interested in exploring with you the needs and ideas for improving the measurement system as it relates to the investigation of the fundamental constants and the underlying physics. To be sure, you can't order up a better volt or a more precisely transferable kilogram for the asking. But some of the toughest problems are applied problems capable of planned attack and solution. For example, when the Josephson effect suggested a way to improve the realization of the volt, the need for an accurate potentiometer serviceable at the nanovolt level was identified and met. This conference can be very useful to us and to other national and international laboratories planning and conducting work on the basic units. It can also serve as a guide to experimentalists everywhere through whose ingenuity new schemes for precision measurement are created. If successful, the conference can not only make it easier to evaluate the constants in the future, but will make a real contribution to the progress of the physics on which better knowledge of the constants depends. For better measurements will result.

The interplay between basic physics, evaluation of the constants and realization of the measurement system not only impacts the accuracy requirements for the measurement system, but the evolution of the definitions or theoretical realizations of the units themselves. Our measurement system is a living language. All physicists should take an active

interest in its evolution, and not leave it to a few measurement experts around the world. How does this language grow?

The course of scientific investigation reveals patterns of systematic behavior and suggests sets of invariant parameters in terms of which observations are most simply described. It will often turn out that within the framework of a specific technique for observation—let's say relative atomic masses or atomic resonances—the precision with which these relative observations can be made exceeds the accuracy with which all the observations can be expressed in units of the basic measurement language. At this point one must re-evaluate the structure of the measurement language. To what extent do we proliferate the creation of new measures—such as the atomic mass unit, the Siegbahn x-unit, the electron volt and the chemical calorie—or do we either change or improve the means for defining, realizing and transferring the primary units of the measurement system?

Philosophically it might be appealing to take the position that one should search for invariant properties of nature, and select from among them those that seem most fundamental to our understanding to be the basic units of measurement. Thus the speed of propagation of electromagnetic waves in free space is a scale factor in our basic geometry. We might like it to be unity. Perhaps next we might assign values to the interaction constants for each of the known classes of forces. But no choice of units will obviate the necessity for the practical interrelation of quantitative observations in all areas of physical phenomenology. First of all, we cannot continue to test the invariance and interrelation of the fundamental constants unless we constantly compare the expectations of theory based on the assumption of their invariance with the results of every physical observation we can make. Assigning numerical values to some of the constants does not eliminate the measurement transfer problem, within which practical gages will always prove both necessary and convenient.

Indeed, the main source of conflict in the evolution of the international measurement language is diversity versus uniformity. At the most sophisticated level, the measurement language not only tolerates but requires a great deal of diversity. Physicists must be multilingual in common measurement speech, in the "natural" systems of units using the fundamental constants as parameters, in SI or cgs electrical units as they please and find convenient, even inventing new units if necessary. Specialists also use short-hand measurement slang, for example: measuring the temperature of a plasma in kilovolts to describe the best fit to a Boltzmann distribution for the excited states of Fe XIV. But as we move to lower levels of sophistication—into applied science, engineering, commerce and finally the general public's use of measurement—the number of people involved grows, the problem of interdisciplinary communication worsens and the need for simplicity and uniformity grows. Since it seems clearly desirable on intellectual grounds for the world's scientists and world public to use the same basic measurement units, and this is indeed the case for the 90 percent of the world's population that talks metric, we must appreciate that we have placed another very important constraint (besides high accuracy) on the basic units. The same theoretical definition must suffice both for the most sophisticated need in science (exploring basic questions of physics) and the most general need of man (dealing with his life and environment in quantitative terms).

Meeting both these needs is the task of the oldest and least conflict-ridden international treaty organization for scientific cooperation the world has known, the General Conference of Weights and Measures established by the Treaty of the Meter, whose centennial we will celebrate in five years. Forty nations adhere to the treaty and meet in full diplomatic formality at regular intervals to consider improvements in the internationally and legally agreed on recommended system of scientific measurement. The General Conference of Weights and Measures conducts its work between conferences through two instrumentalities: the Comité International des Poids et Mesures (CIPM) with its several consultative committees and a unique international laboratory, the Bureau International des Poids et Mesures (BIPM) at the Pavillon de Breteuille in Sèvres near Paris. The Director of this laboratory is of course our Chairman, Dr. Terrien. Any of you who have not visited the BIPM should do so. As you enter the gate you will leave France and enter international territory, but I assure you that Dr. Terrien's constabulary are quite hospitable to international travellers—especially if they are scientists.

Through the work of this laboratory, and the Consultative Committees of the CIPM—particularly that dealing with Units—the General Conference has developed the International System of Units (SI). National governments and other international organizations—both official such as the International Organization for Legal Metrology (OIML) and private like the International Organization for Standardization (ISO)—are taking actions that will steadily increase the universality of SI units in commercial and industrial usage. Whether scientists like to speak SI or not, they will find that increasingly the industrially manufactured equipment they use will be displaying output in SI units.

In view of the fact that for scientific usage, SI is entirely voluntary—at least in this country—a surprising amount of emotional energy has been invested in its development. I could extend these remarks by another 30 minutes just to digress into the question "What is a Base Unit?" and "If so, why is the candela one of them?" I am sure you will be fascinated to know that a good many technical manhours have been expended arguing the question, "Should the dimensionless numbers form one of the Base Units?" I could also digress into the issue of the consequences to the U.S. of being the last industrialized nation on earth to face up to the question of

conversion to general use of SI units instead of the inch, yard, fathom, chain, rod and furlong. NBS is engaged in a major study for the Congress on this question, with a report due in 12 months. But for the moment I wish only to entreat you to take an active interest in the question of the redefinition of the basic physical units, as well as in the improvement in effective accuracy of the units as now defined. NBS people (along with the Naval Observatory) play a major part in behalf of the U.S. in the painstaking work and sometimes frustrating negotiations that matters like redefinition of the meter or second involve. We want help from our university and industry colleagues; we have no desire to monopolize the expression of U.S. views on such important scientific matters. I might also add that Dr. Terrien is also grateful for help from interested scientists everywhere—including the recommendation to him of gifted younger scientists who might wish to apply for appointments on the BIPM staff. Since BIPM is an international laboratory it is our duty to respond.

Now let me turn to the matter often referred to as the "adjustment of fundamental constants." This phrase has always bothered me, and I prefer the phrase "evaluation." Adjustment of constants connotes a physicist with a screwdriver diddling the parts-per-million dial of some cosmic odometer. Having just returned from a 1200 mile sailing voyage on open ocean with a binnacle compass suffering from an accidentally "adjusted" compensating magnet I don't think God meant either man's compasses or His constants to be "adjusted" frivolously.

Let me start by emphasizing once more that the purpose of this conference is NOT to produce another evaluation of the constants. The Conference announcement says,

> "The objective of the conference is to discuss modern techniques of precision physical measurement and their application, along with modern theoretical developments, to the determination of the fundamental constants. It is not the purpose of the meeting to arrive at new adjusted values of the constants, but it is expected that the substantive discussions at the conference will help provide a basis for future adjustments."

Furthermore, it is not the intent of the organizers to restrict in any way anyone's freedom to undertake such an evaluation if he wishes. Quite the contrary. Careful evaluation of the constants has been a labor of love for many years. For the devotion of a few dedicated people all of science should be grateful. Starting with the International Critical Tables and the Handbuch der Physik in 1926, the history of the most notable evaluations—at least in this country—surrounds only a few names.

As Editor of the *Reviews of Modern Physics* it is a proud fact that page one of volume one in 1929 contained R. T. Birge's first review [5], to be followed in 1932, 1942 and 1945. An equal span of years embraces the contributions of Jesse DuMond and E. R. Cohen [6]. Important reviews have been provided by Bearden, Earle, Minkowski and Thomsen [7] and most recently by Taylor, Parker and Langenberg [4]. The last mentioned of these reviews will serve to illustrate my own view on the issue of uniformity versus diversity with respect to the evaluation of the constants and recommendations of best values.

Many purposes may motivate an evaluation of the constants, but two limiting cases must be recognized. Since the constants are the parameters for theoretical physics, just as the units are the language for the operational description of observations, every scientist using physics theory must use values of the constants when comparing theoretical expectation with reality. In the overwhelming majority of the cases the fully available precision of our knowledge of the constants is not needed. But since the user can scale his prediction with the values of the constants used as parameters it is absolutely essential that he know and accurately communicate what values he actually used. Obviously this problem is simplified if everyone in the world uses the same values of the fundamental constants. Short of assigning these constants arbitrary values, for example unity, this requires agreement on what values to use and some means for changing them as knowledge improves. On the other hand, the process of evaluation is itself a matter of physics judgment. Indeed it is one of the most demanding tasks a physicist can face, for he must reliably extract all the information content in an enormous range of experiments in the context of very sophisticated theoretical understanding. Thus there can be no unique "best" values that are not arrived at either by autocratic methods or methods that throw away information. Yet it is only from the use of all the information that the consistency of our understanding of physics can be tested. How can both the general user of physical constants and the specialist probing the limits to which theory has been shown to be reliable be served?

We must have *both* active intellectual competition in the evaluation of the constants, and a reliable, effective means for worldwide adoption of recommended values.

For the first purpose, *Reviews of Modern Physics* was pleased to receive the paper of Taylor, Parker, and Langenberg [4] precisely because it was a thoughtful but still personal presentation of their views as they sought to obtain best values—best in the sense that they best suited their own purpose in making the adjustment. And they had the courage to reach a conclusion, namely that "it appears that all is not well with quantum electrodynamics." That may or may not be the consensus view of those who are here—indeed the consensus might be different on Thursday from what it is today. But I hope it stimulates others to consider the problem from alternative viewpoints and do an equally conscientious job. I suspect that somewhat different constants might result; I am not sure the error

statements would overlap. But even more important, different conclusions about physics might emerge.

Yet if we were to receive a series of such reviews—say one every year, chaos might result from the confusion of literature citations, and the difficulty of finding out what constants were put into each of thousands of computer programs in order to tell whether the data or the algorithms caused identical theories to yield numerically different results.

The recommended cure: CODATA, the Committee on Data for Science and Technology of the International Council of Scientific Unions. CODATA's current president is F. D. Rossini of the University of Notre Dame, who retires from that post in September. CODATA has a permanent Bureau in Frankfort, Germany, and a Task Group on Fundamental Constants, chaired by E. R. Cohen. This Task Group now faces the problem of deciding how to proceed.

For the U.S. nationals present, let me note that we participate in the CODATA Task Group through the Committee on Fundamental Constants of the Numerical Data Advisory Board. This Board is established by the two Academies: Science and Engineering. Chairman of the U.S. Committee is J. S. Thomsen; other members are Bender, Cohen, Deslattes, Erickson, Ramsey, and Taylor. There is a U.S. National Committee for CODATA—also part of the activities of the NDAB. Robert Brode is Chairman of it; I am the U.S. Delegate to CODATA. So much for bureaucracy. The important problem is—what should be the strategy of the CODATA Task Force?

I suggest that CODATA should not attempt to publish a set of "best" values—which cannot be competently done by a committee, especially an international committee. They should publish—preferably at regularly scheduled intervals—a set of "recommended" values to be used by those whose needs are satisfied by less than the full accuracy that is potentially available. Those who want to fight about a fourth order correction in QED, or who suspect that space might not be isotropic would be referred to their friendly literature.

This still leaves unanswered just what numerical values to use and how to describe the errors. I hope many of you will join in the round table on Friday afternoon to explore different views on how the constants should be combined statistically. But hopefully the errors would be conservative and every major standards laboratory would adopt the values recommended for routine use in their own work, reviews, and other information systems they operate. Indeed an international, multilingual set of pocket cards, whose color changes with each updating, might be a good project for CODATA if its Task Force is successful in resolving these issues.

Finally, let me go back to an earlier remark about "meaningful" comparisons between experiment and theory. My brief experience as editor of a journal that is somewhat grimly insisting that authors *critically evaluate* the data with which they compare theoretical expectations is very disappointing. It is obvious that many scientists simply do not—or cannot—do this. Sometimes the reason is laziness, or lack of experience in experimental physics. But all too often it is because the literature they are reviewing simply cannot be critically evaluated because the information needed for the evaluation is not there [8, 9].

If any of you have not read the "Recommendations for Reporting Experimental Results" in the final pages of the Taylor, Parker, and Langenberg review, I urge you to do so. They sum it up in one sentence [10]: "The experimenter should ask himself, Have I provided enough information in a sufficiently clear manner so that ten years hence my result can be updated in light of any new information or data which may become available?" For if a paper meets that test, it will also meet the test of suitability for evaluation, which is the price of a finite weight factor in any statistical averaging with other data of like kind.

This is a particularly propitious time for this conference. Discovery of the laser and the Josephson effect have each had their impact on precision measurement and are well on the way to impacting the definitions and means for transferring the basic units themselves. The CODATA Task Force is recently organized as the first official international body to address itself to the question of recommending values for the physical constants. And while the rest of physics divides itself into squabbling subcultures of over-specialization, we draw knowledge from many corners of the physical world, synthesizing a view of nature's basic constancy, never knowing when another layer of phenomenology may be revealed. If you are sustained by the possibility—just the bare possibility—that c varies with time or orientation, if you can persuade yourself that the ratio of flux quantum to chemical potential in a superconducting junction might depart just a teensy bit from 2, it is not so hard to take being classified as "General Physics" or even "miscellaneous" in the bureaucracy of grant-giving agencies.

It is a pleasure for me to open this International Conference on Precision Measurement and Fundamental Constants. A world obsessed with the pervasive effect of rapid change might take some solace from this conference. For while mankind has abused nature—with industrial, agricultural and human destruction and contamination and may find nature fickle, or even hostile as a result—physicists who search for her truths with delicate precision find her steadfast and true. We have learned that the deeper we wish to probe, the more gentle we must be. By listening carefully, looking closely and touching ever so gently, we can slowly peel away the layers of our ignorance. Our faith in the rationality of the natural order is reinforced; our concern for mankind's irrationality heightened. Ours is not a clumsy or utilitarian art, but the highest refinement of scientific skill in both its experimental and theoretical aspects. By asking the most fundamental questions of science we generate and refine a system of practically realizable measurement units and tools

in terms of which to probe all physical phenomena. This continually evolving system then provides the basis for the development of all scientific knowledge, without which we cannot live in harmony with nature.

References

[1] *Proceedings of the First International Conference on Atomic Physics*, Bederson, B., Cohen, V. W., and Pichanick, F. M. J., Editors (Plenum Press, Inc., New York, 1969).

[2] *Physics of the One- and Two-Electron Atoms*, Bopp, F., and Kleinpoppen, H., Editors (North-Holland Publishing Co., Amsterdam, 1969).

[3] Nuovo Cimento Suppl. **6** (1957).

[4] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[5] Birge, R. T., Rev. Mod. Phys. **1**, 1 (1929).

[6] See for example, Cohen, E. R., and DuMond, J. W. M., Rev. Mod. Phys. **37**, 537 (1965).

[7] See for example, Bearden, J. A., and Thomsen, J. S., Nuovo Cimento Suppl. **5**, 267 (1957).

[8] Branscomb, L. M., The Misinformation Explosion, Scientific Research, May 27, 1968 (McGraw-Hill, Inc.).

[9] Branscomb, L. M. Truth in Packaging Scientific Information, Measurements and Data **3** No. 5, 104 (September–October 1969).

[10] Loc. cit., p. 485.

My old prof said (it has not been regretted)
"Here's a rule that should ne'er be forgetted:
Evaluate, be just,
Adjust, if you must,
And let constants, but not data, be fitted."

DONALD HALFORD

FREQUENCY AND TIME STANDARDS

Standards de Frequence

C. Audoin and M. Arditi

Laboratoire de l'Horloge Atomique, C.N.R.S. Bâtiment 220, Faculté des Sciences, 91 Orsay, France

This survey reviews the fundamental physics of the more advanced types of primary atomic frequency standards (cesium beam frequency standard and hydrogen maser). Performances and accuracy limiting factors are discussed. Possible future developments are outlined.

Key words: Cesium beam frequency standard; frequency standard; hydrogen maser.

1. Introduction

1.1. Historique

Dans les années 1940, les standards de fréquence étaient des oscillateurs à quartz. Leur fréquence est définie par une résonance mécanique. Elle possède l'inconvénient grave de présenter un effet de vieillissement et d'être très sensible aux conditions extérieures.

Ces difficultés furent surmontées dans les années 1950 en utilisant une résonance atomique pour contrôler la fréquence de l'oscillateur à quartz par l'intermédiaire d'une boucle d'asservissement de fréquence. Les effets de vieillissement disparaissent et ceux de l'environnement sont considérablement réduits.

L'étude extensive des propriétés et des performances des étalons de fréquence atomique com-

TABLEAU 1. *Frequence de transition hyperfine de l'hydrogène et de ses isotopes, relativement à celle du césium 133*

Elément	Spin nucléaire	Fréquence de transition hyperfine (Hz)	Incertitude de mesure
Cesium	7/2	9 192 631 770	Définition de la seconde
Hydrogène	1/2	1 420 405 751,768	2×10^{-12}
Deutérium	1	327 384 352,51	$1,5\times10^{-10}$
Tritium	1/2	1 516 701 470,7919	$4,7\times10^{-12}$

prenant un résonateur à jet de césium, ainsi que la grande diffusion de cet appareil a conduit la 13e Conférence Générale des Poids et Mesures, le 13 Octobre 1967, à décider que l'unité de temps du Système International d'Unités est la seconde définie dans les termes suivants:

"La seconde est la durée de 9.192.631.770 périodes de la radiation correspondant à la transition entre les deux niveaux hyperfins de l'état fondamental de l'atome de césium 133".

Dans les années 1960, les masers, et plus spécialement les masers à hydrogène ont été développés suffisamment pour pouvoir être utilisés comme standards primaires de fréquence. Ici, l'énergie nécessaire à l'entretien de l'auto-oscillation dans une cavité résonante est fournie par les atomes eux-mêmes. Un asservissement de phase assure le contrôle de la fréquence de l'oscillateur à quartz par celle du maser.

Il apparaît dès maintenant que les années 1970 seront marquées par les progrès réalisés dans la stabilisation de fréquence des lasers et par la possibilité de comparer les fréquences optiques et radioélectriques.

1.2. Précision des Mesures de Fréquence

Les étalons atomiques de fréquence [1] sont, comme l'étalon actuel de longueur (lampe à krypton 86), fondés sur l'observation des propriétés internes des atomes.

Tous les atomes d'une même espèce atomique ont, selon un postulat de base de la physique atomique, les mêmes propriétés en tout point et à tout instant, à des effets relativistes près (ceux-ci ne seront pas discutés ici).

Ces étalons primaires peuvent donc être construits en de nombreux exemplaires, et mis à la disposition de nombreux expérimentateurs. Cette propriété les distingue fondamentalement de l'étalon de masse par exemple.

Actuellement, les mesures de fréquence, et plus particulièrement celles des fréquences de transition hyperfine du césium, de l'hydrogène [2] et de ses isotopes [3, 4] à l'état fondamental sont les mesures physiques les plus précises (tableau 1). Par exemple, l'incertitude de la mesure de la fréquence de transition hyperfine de l'hydrogène, relativement à celle du césium 133 est de $\pm2\times10^{-12}$.

Cette précision doit être attribuée à deux facteurs:

(a) les principes physiques mis en jeu dans les standards de fréquence permettent d'obtenir des raies spectrales très étroites d'une part, et de les observer dans des conditions où l'influence des facteurs externes (champs magnétiques et électriques, température, etc) est très faible, d'autre part.

(b) le développement de la technique radioélectronique, et en particulier l'utilisation de la non-linéarité de certains composants, permet des comparaisons précises entre fréquences différentes, par l'intermédiaire d'opérations arithmétique exactes sur les fréquences: somme, différence, multiplication et division. Ces opérations (à l'exception de la division) sont maintenant effectuées couramment jusque dans le domaine infrarouge [5, 6] où la non linéarité de matériaux optiques peut également être utilisée.

2. Description des Standards de Fréquence à Césium et à Hydrogène

Nous limiterons notre exposé à la description des étalons de fréquence primaires. Les horloges atomiques à cellule de rubidium pompé optiquement, dont la fréquence dépend de la pression de gaz tampon ne seront pas traitées, malgré leur excellente stabilité de fréquence à court terme et leur intérêt en tant qu'étalon de tranfert.

A l'heure actuelle, parmi les étalons primaires de fréquence, seul l'étalon de fréquence à jet de césium est disponible commercialement.

2.1. Choix d'une Transition

Un étalon de fréquence idéal serait celui dans lequel l'atome serait au repos dans l'espace libre, la transition utile étant observée pendant un temps très long. Cet étalon n'est pas réalisable en pratique.

Cependant, on peut s'approcher de ces conditions idéales. Pour cela, on choisit convenablement la transition, ainsi que le manière dont elle est observée. Lorsqu'il en est ainsi, c'est-à-dire lorsque l'écart de la fréquence de la transition, par rapport au cas idéal, est faible et calculable—ou mieux mesurable—avec un haut degré de précision, l'étalon est dit primaire.

A l'heure actuelle, les étalons primaires de fréquence les plus répandus sont l'appareil à jet de césium et le maser à hydrogène. Ils permettent la mesure de la fréquence de transition hyperfine $\Delta F=1$, $\Delta m=0$ des éléments correspondants.[1]

Les propriétés intéressantes de ces transitions dipolaires magnétiques, caractéristiques des atomes ou des ions possédant un seul électron périphérique sont les suivantes:

—la fréquence de transition n'est déplacée qu'au

[1] L'emploi de la structure hyperfine d'autres éléments a été également envisagé. Ces éléments sont: ^{21}Na, ^{39}K, ^{85}Rb, ^{87}Rb, ^{205}Tl. Il en est de même pour les transitions dipolaires électriques de H_2S, CO, SO_2, BaO, LiF, $^{14}NH_3$, $^{15}NH_3$, HCN.

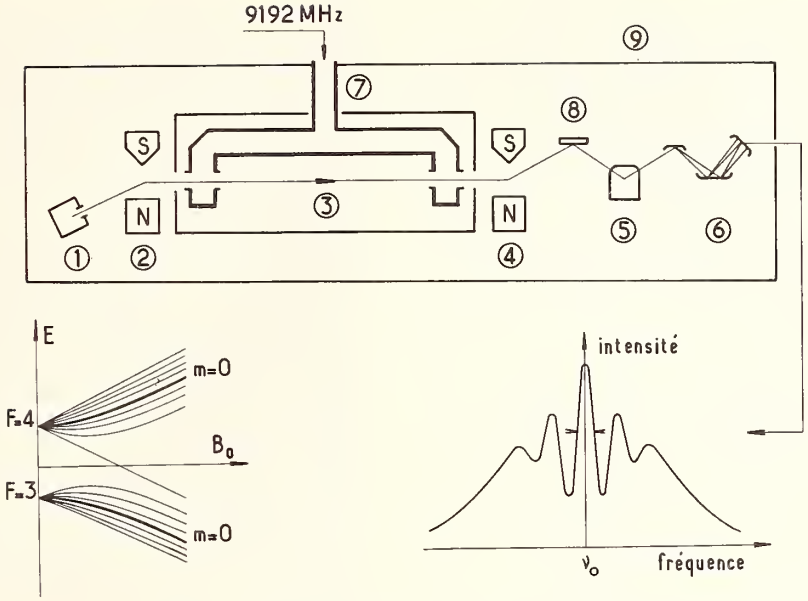


FIGURE 1. *Schéma de principe du résonateur à césium.*
En haut: 1. Four à césium—2. Aimant de sélection—3. Région blindée magnétiquement où l'on produit un champ magnétique continu de 50 mG environ—4. Aimant d'analyse—5. Spectro de masse—6. Multiplicateur de particules—7. Cavité résonnante à deux bras—8. Ioniseur à fil chaud—9. Enveloppe étanche.
En bas: Structure hyperfine du césium et réponse du résonateur

second ordre par le champ magnétique directeur. Par example, dans des conditions typiques, le déplacement relatif, calculé à partir de la formule de Breit-Rabi est de 2×10^{-12} pour l'hydrogène à 1 mG et de $1{,}1\times10^{-10}$ pour le césium à 50 mG, valeurs de champ typiquement utilisées dans les appareils courants.

Ce déplacement est mesurable avec une grande précision, meilleure que le millième, par l'intermédiaire des transitions $\Delta F=0$, $\Delta m=\pm1$ dont la fréquence est proportionnelle à la valeur du champ magnétique.

—le déplacement de fréquence par effet Stark est fonction quadratique du champ électrique appliqué. Il est parfaitement négligeable dans les conditions habituelles.

—le couplage entre le moment magnétique de l'électron périphérique et celui du noyau étant faible, les fréquence de transition appartiennent au domaine micro-onde.

2.2. Nécessité de la Séparation des États de Structure Hyperfine

La différence d'énergie $h\nu$ entre les deux niveaux atomiques est, pour ces transitions de micro-onde, petite devant kT. (La situation est inverse dans le domaine optique.) Par exemple, $h\nu/kT=1{,}6\times10^{-3}$ pour $\nu=10$ GHz et $T=300$ K. Selon la loi de Boltzman, il en résulte que les niveaux entre lesquels a lieu la transition sont sensiblement de populations égales.

Pour déceler la transition avec un rapport signal/bruit convenable, il est donc nécessaire de rompre l'équilibre thermodynamique.

Pour le césium et l'hydrogène, la séparation des états est magnétique. Elle s'effectue sur un jet, comme dans l'expérience de Stern et Gerlach et généralement, seuls les atomes appartenant à l'un des deux états considérés sont dirigés vers la région d'interaction avec le champ électromagnétique.

2.3. Le Résonateur à Jet de Césium

Le schéma de principe du résonateur à césium, dans sa version la plus courante, est représenté sur la figure 1.

Un four suivi d'un collimateur produit un jet d'atomes de césium dans une enceinte où règne un vide poussé.

Un premier aimant, qui produit un champ magnétique intense (de l'ordre de 1 Tesla) fortement inhomogène, effectue la sélection d'états. Les atomes d'état $F=4$, $m_F=0$ forment alors un faisceau parallèle à l'axe de l'appareil. Les atomes d'état $F=3$, $m_F=0$ frappent les parois où ils se condensent.

Le faisceau traverse ensuite la région d'interaction avec le champ électromagnétique. Celui-ci est appliqué en deux régions distinctes, à l'extrémité des deux bras d'une cavité résonante selon une disposition proposée par Ramsey [7]. Un blindage magnétique entoure la cavité; à l'intérieur de celui-ci, un champ magnétique homogène de 50 mG environ est produit par une bobine. Un second aimant à champ inhomogène analyse enfin le faisceau. Dans cette version, il dirige vers le détecteur les atomes qui ont subi la transition 4,0→3,0 par émission stimulée. Habituellement, un spectromètre de masse et un multiplicateur de particules suivent le détecteur à ionisation de surface pour éliminer les ions parasites et améliorer le rapport signal/bruit du signal de sortie. Celui-ci

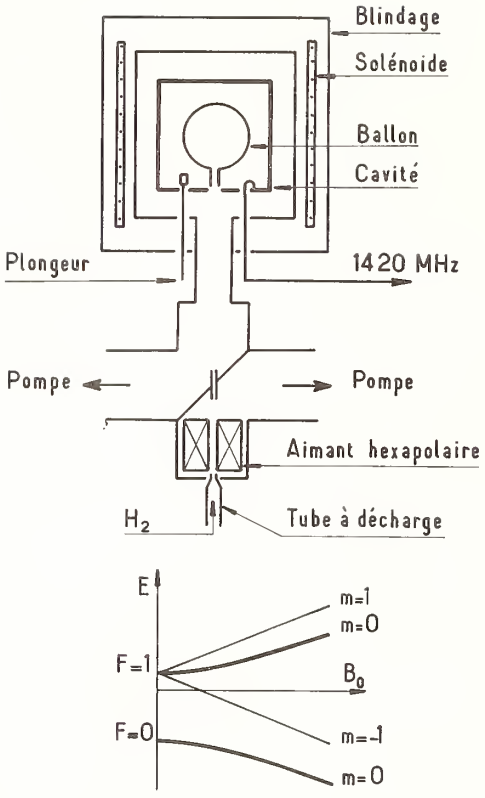


FIGURE 2. *En haut: Schéma de principe du maser à hydrogène. En bas: Structure hyperfine de l'atome d'hydrogène.*

présente un maximum à la fréquence hyperfine du césium.

Une propriété importante de l'appareil est de permettre l'observation de la raie de résonance non élargie par effet Doppler. Seul subsiste l'effet Doppler transverse du second ordre qui déplace la fréquence de la quantité $\Delta f/f_0$, en valeur relative, donnée par:

$$\Delta f/f_0 = -\tfrac{1}{2}(v^2/c^2)$$

où v est la vitesse des atomes et c la vitesse de la lumière. Pour $v = 230$ m/s, $\Delta f/f_0 = -3\times10^{-13}$.

La résonance est donc observée avec sa largeur naturelle qui est inversement proportionnelle au temps d'interaction avec le champ électromagnétique, en conséquence de la relation d'incertitude de Heisenberg. Avec un résonateur de 55 cm de long, la largeur de la raie de résonance est de 300 Hz environ.

2.4. Le Maser à Hydrogène

L'auto-oseillation d'un maser (Microwave Amplification by Stimulated Emission of Radiation) a été obtenue, pour la première fois, avec l'ammoniac, en utilisant une transition de son spectre d'inversion. De meilleurs résultats, dans l'emploi en étalon de fréquence sont obtenus avec l'hydrogène atomique dont le spectre de structure hyperfine est le plus simple que l'on puisse envisager, le spin nucléaire et le spin électronique étant égaux à 1/2. Cependant, la faible probabilité des transitions dipolaires magnétiques, comparée à celle des transitions dipolaires électriques requiert l'emploi d'une cellule de stockage des atomes dans la cavité résonnante pour obtenir un temps d'interaction suffisant avec le champ électromagnétique et obtenir l'auto-oscillation. Le schéma de principe du maser à hydrogène est indiqué sur la figure 2 [8, 9].

La transition utilisée est la transition de structure hyperfine $F=1$, $m_F=0 \rightarrow F=0$, $m_F=0$ de l'atome d'hydrogène à l'état fondamental.

L'hydrogène atomique est produit par dissociation de l'hydrogène moléculaire dans une décharge haute-fréquence. Il émerge de la source à travers un collimateur et forme un jet atomique dans un récipient où le vide est entretenu par des pompes. Les atomes passent ensuite dans un sélecteur d'état constitué par un aimant hexapolaire produisant un champ magnétique très inhomogène. Il se comporte comme une lentille qui focalise les atomes d'états $F=1$, $m_F=0$ et $F=1$, $m_F=1$ qui seuls pénètrent dans le ballon de stockage. Les atomes d'état $F=1$, $m_F=1$ sont inutiles à la transition d'horloge. Ils peuvent être éliminés par passage adiabatique rapide [10].

Le ballon de stockage, en quartz, est placé au centre d'une cavité résonnante cylindrique; il est revêtu intérieurement de téflon, un matériau inerte et non désorientant pour les atomes d'hydrogène. Les atomes quittent le ballon après y avoir séjourné 1 seconde environ.

Les atomes sont maintenus, avec une vitesse d'ensemble sensiblement nulle dans un ventre du champ électromagnétique de la cavité. Il en résulte que l'effet Doppler du 1er ordre est éliminé. Seul subiste l'effet Doppler du second ordre. Les atomes étant en équilibre thermique avec la paroi de la cellule, le déplacement relatif $\Delta f/f_0$ a pour expression

$$\Delta f/f_0 = -3kT/2mc^2$$

où $3kT/2$ est l'énergie d'agitation thermique des atomes, m est leur masse et c la vitesse de la lumière. A 300 K et pour l'hydrogène atomique, $\Delta f/f_0 = 4{,}3\times10^{-11}$. Ce déplacement est, en principe, mesurable avec la même précision que la température de la cellule, c'est-à-dire à mieux que 1/1000 près.

Ici, également, la résonance est observée avec sa largeur naturelle. Celle-ci est de l'ordre de 1 Hz, correspondant à un coefficient de surtension de 10^9, valeur favorable à l'obtention d'une très grande stabilité de fréquence. Cette surtension est égalée dans une cellule à méthane, utilisée pour le contrôle de la fréquence d'un laser He–Ne oscillant à 3,39 μm, par l'artifice de l'absorption saturée [11]. Une telle sélectivité n'est dépassée que par l'emploi de cavités électromagnétiques à parois supra-conductrices d'une part, et par l'affinement des spectres de rayonnement γ par effet Mössbauer d'autre part.

La région d'interaction est soigneusement blindée magnétiquement. Un champ magnétique uniforme, de 1 mG environ, produit par un solénoïde est appliqué aux atomes.

Lorsque l'énergie apportée à la cavité par le jet dépasse les pertes de diverses origines (pertes par

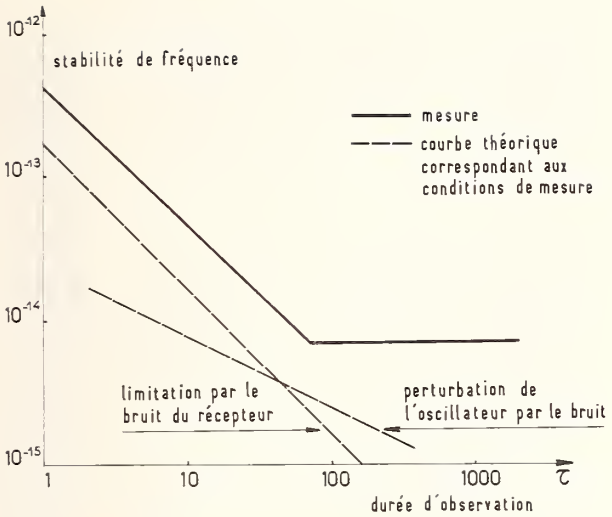


FIGURE 3. *Stabilité de fréquence du maser à hydrogène.*

effet Joule dans les parois de la cavité, relaxation), le maser oscille à la fréquence de 1420 MHz. La puissance disponible est de l'ordre de 10^{-13} W.

La stabilité de fréquence de cet oscillateur est, en fait, limitée par le bruit thermique des circuits de mesure en raison de la faible puissance d'oscillation. Malgré cela, le maser à hydrogène est à l'heure actuelle, l'oscillateur qui possède la meilleure stabilité de fréquence, pour les temps de mesure supérieurs à 10 s [12]. Celle-ci est indiquée sur la figure 3.

2.5. Contrôle d'un Oscillateur à Quartz

Les fréquences des transitions de structure hyperfine appartiennent à un domaine facilement accessible à la technique radio-électrique courante. Ces transitions sont, pour les besoins pratiques, utilisées pour contrôler la fréquence d'un oscillateur à quartz.

La présence d'un oscillateur auxiliaire est, par principe, nécessaire pour détecter la résonance dans un dispositif passif comme le résonateur à jet de césium. Ce n'est pas le cas pour un auto-oscillateur qui, comme le maser à hydrogène, fournit son propre signal qui peut alimenter un amplificateur à bande étroite. Cependant, un générateur auxiliaire est souvent utilisé, par raison de commodité, pour disposer d'un signal dont la fréquence a une valeur normalisée, 5 MHz par exemple.

Dans les deux cas, l'oscillateur à controler doit avoir une excellente stabilité à court terme car l'effet de l'asservissement n'est sensible que pour des temps de mesure supérieurs à une valeur qui est caractéristique des propriétés de l'asservissement. On utilise toujours un oscillateur à quartz pour cette application.

(a) *Cas du Résonateur Atomique.* Le schéma de la figure 4a montre, sous forme très condensée, le principe de l'asservissement d'un générateur sur la fréquence d'une résonance atomique. La fréquence f_g du générateur est modulée de façon à balayer le sommet de la courbe de résonance. Le signal obtenu est analysé dans un démodulateur pour détecter l'écart de la fréquence f_g par rapport à la fréquence de résonance f_0. Un signal d'erreur dépendant de l'écart $f_g - f_0$ est appliqué au générateur. Finalement, sa fréquence est asservie à celle de la résonance.

Le même type de montage se retrouve dans la stabilisation de fréquence d'un laser sur une résonance moléculaire.

Dans le cas de l'horloge à césium, le générateur auxiliaire comprend un oscillateur à quartz suivi d'un synthétiseur de fréquence.

Des précautions doivent être prises pour assurer une mesure correcte de la fréquence de résonance. Il faut veiller à ce que le signal d'excitation ne perturbe pas la résonance atomique, par la présence de signaux de fréquence voisine par exemple. De plus, la modulation de fréquence doit être bien symétrique autour de la fréquence du générateur. La technique actuelle permet d'éviter, de façon satisfaisante, ces sources d'erreur systématique.

La stabilité de fréquence de l'oscillateur asservi est limitée par le bruit de grenaille du jet atomique [13]. Dans les appareils courants, l'écart type des

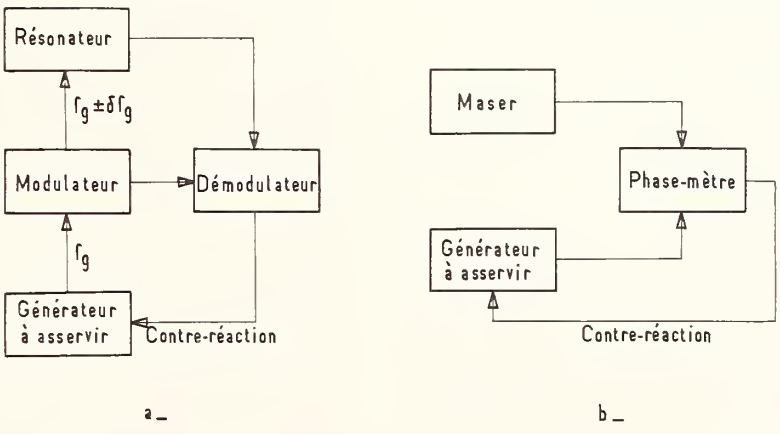


FIGURE 4. a—*Principe de l'asservissement d'un générateur sur un résonateur.* b—*Principe de l'asservissement d'un générateur sur un oscillateur.*

fluctuations relatives de fréquence est de l'ordre de:

$$\sigma(\Delta f/f_0)\approx 6\times 10^{-11}/\tau^{1/2}$$

où τ est le temps de mesure, exprimé en seconde.

Pour $\tau > 10^5$ s, la stabilité de fréquence est limitée par des causes de fluctuations supplémentaires. Elle est cependant très bonne, de l'ordre de 4×10^{-14} pour 90 jours, et de 10^{-13} par an dans les appareils commerciaux.

(b) *Cas du Maser.* Le principe de l'asservissement est indiqué sur la figure 4b. Les signaux provenant du maser et du générateur sont comparés dans un phasemètre dont l'indication est utilisée pour contrôler le générateur. Celui-ci est bloqué en phase sur l'oscillation du maser.

Contrairement au cas précédent, l'asservissement ne risque pas d'introduire une erreur systématique dans la mesure de la fréquence de la résonance atomique.

Le générateur comprend un oscillateur à quartz et des circuits de synthèse de fréquence.

3. Comparison des Standards. Limitations—Exactitude

La figure 5 permet la comparaison des appareils à jet, ou à jet et ballon de stockage qui ont été utilisés pour des mesures de fréquence de transition hyperfine.

3.1. Résonateur à Jet

Historiquement, l'ancêtre des standards de fréquence est l'appareil à jet mis au point par Rabi et ses collaborateurs [14] pour la mesure des moments magnétiques nucléaires et pour des études de spectroscopie (fig. 5a). Les aimants de séparation et d'analyse sont semblables à ceux que l'on rencontre le plus souvent dans les horloges à césium. Par contre, le champ électromagnétique est appliqué aux atomes tout le long de la structure résonnante, en forme d'épingle à cheveux (hairpin).

Une amélioration due à Ramsey [7] a consisté à appliquer le rayonnement en deux régions séparées (fig. 5b). Cette technique permet une réduction de 40% de la largeur de raie, si l'on se réfère au cas où le rayonnement est appliqué sur tout l'intervalle séparant ces deux régions.

Cette disposition présente en outre des avantages dans la réduction de l'effet des inhomogénéïtés du champ directeur et dans l'élimination de l'effet Doppler du premier ordre.

Les inévitables défauts de symétrie entre les deux bras de la cavité résonnante provoquent une distorsion de la raie se traduisant par une erreur sur la mesure de la fréquence de résonance. C'est là, le principal facteur limitant l'exactitude[2] de ce type d'horloge atomique [15]. A un degré moindre, interviennent l'effet des inhomogénéïtés du champ directeur et l'incertitude sur la valeur de l'effet Doppler

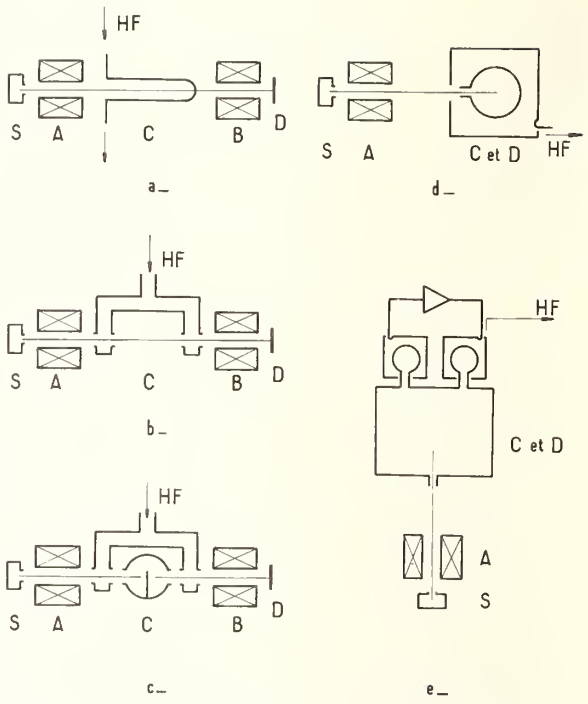


FIGURE 5. *Historique et comparaison des standards de fréquence.*
a—Appareil de Rabi
b—Appareil à deux excitations séparées
c—Appareil à deux excitations séparées et stockage
d—Maser à hydrogène classique
e—Maser à hydrogène à très grand ballon de stockage
S: source—A: aimant de séparation—B: Aimant d'analyse—C: région d'interaction—D: détecteur—HF: signal haute fréquence

du second ordre. L'exactitude des étalons de fréquence à césium du commerce est très bonne, de l'ordre de $\pm 5\times 10^{-12}$.

L'effet du déphasage dans la cavité peut être mesuré, dans les étalons de laboratoire, en inversant le sens du jet, relativement au résonateur, ou en modifiant la vitesse moyenne des atomes. Leur exactitude est comprise entre 5×10^{-13} et 10^{-12}. Il semble possible d'atteindre la valeur de 10^{-13}.

La stabilité à court terme étant proportionnelle, à longueur de cavité constante, à $I^{-1/2}$ où I est le flux atomique atteignant le détecteur, des études sont menées pour augmenter ce flux. Pour cela, on optimise les performances des dispositifs à aimants dipolaires, les plus couramment employés, ou on utilise des aimants de focalisation hexapolaires ou quadrupolaires [16, 17]. Cependant le détecteur est de grande surface, le bruit de fond augmente, et les performances ne sont pas améliorées. Récemment, une nouvelle disposition a été proposée pour améliorer le rapport signal/bruit à la détection [16].

Par ailleurs, on peut combiner les techniques des jets atomiques avec celles du pompage optique. On peut en effet remplacer les sélecteurs et analyseurs magnétiques par des dispositifs optiques producteurs de dissymétrie et détecteurs de dissymétrie [18]; la résonance atomique étant toujours provoquée dans une cavité résonnante à deux bras, où la lumière de pompage ne pénètre pas.

L'expérience a été réalisée récemment sur un jet d'atomes de ^{87}Rb [19].

[2] L'exactitude est, par convention, numériquement égale à l'incertitude sur la valeur mesurée de la fréquence de résonance de l'atome supposé isolé dans l'espace vide.

Il est possible d'utiliser des faisceaux atomiques larges et denses, ce qui augmente la sensibilité de la détection et simplifie l'appareil. Grâce à la suppression des champs magnétiques intenses non uniformes, il est possible d'appliquer tout le long du trajet des atomes un champ magnétique uniforme et de faire la résonance magnétique dans des conditions parfaitement définies.

En principe, la technique du pompage optique devrait permettre d'obtenir des raies plus fines que les méthodes de déviation magnétique. En effet, si l'on travaille dans la région où les effets sont linéaires, ce qui est pratiquement le cas avec les intensités des sources lumineuses dont on dispose, l'efficacité du pompage optique et la grandeur du signal de détection sont l'une et l'autre proportionnelle, au temps de séjour des atomes dans la région éclairée, ce qui favorise les atomes lents du jet. Le calcul montre qu'on peut gagner un facteur de l'ordre de 1,4 en finesse de raie par rapport aux méthodes de déviation magnétique. Des expériences en cours semblent vérifier ces prédictions.

Au point de vue précision, ce standard de fréquence est probablement soumis aux mêmes limitations fondamentales que le standard à jet de césium classique. Sa conception mécanique plus simple pourra peut-être lui être favorable dans certains cas.

3.2. Appareil à Jet et Cellule de Stockage

On a cherché à réduire la largeur de la raie de résonance pour réduire l'effet d'un déphasage résiduel entre les deux bras de la cavité et pour diminuer l'erreur sur le pointé du sommet de la courbe de résonance par le système électronique de mesure. Deux voies ont été suivies:

—des cavités très longues ont été réalisées (jusqu'à 4 m), mais il en est résulté une perte de l'intensité du jet et de la stabilité à court terme, ainsi qu'une augmentation de l'incertitude sur l'effet du champ magnétique directeur à cause de la difficulté de produire un champ sufisamment homogène sur une grande longueur.

—les atomes ont été stockés dans une cellule enduite de téflon ou de paraffine située entre les deux bras de la cavité [20] (fig. 5c). Ils devaient alors effectuer plusieurs chocs sur la paroi avant de sortir du ballon pour traverser la deuxième région d'interaction. L'expérience a montré que le nombre de chocs restait faible, limité à 200 environ et que des déplacements de fréquence apparaissaient.

3.3. Maser à Hydrogène

Cette idée du stockage a été reprise avec succès dans le maser à hydrogène, à cause de la plus faible polarisabilité de l'hydrogène atomique, par comparaison au césium.

L'interaction avec le rayonnement se produit dans une seule cavité, comme dans la méthode de Rabi, mais elle a lieu dans une cellule (fig. 5d). Le temps d'interaction devient très long, l'effet maser devient possible dans une cavité à faibles pertes.

L'effet du déphasage entre les deux bras de la cavité est évidemment supprimé, mais il est remplacé par une dépendance de la fréquence d'oscillation vis à vis de la fréquence d'accord de la cavité (cavity pulling). L'effet de l'entraînement de fréquence par la cavité peut être maîtrisé par une régulation thermique des dimensions de la cavité. Les dérives lentes résiduelles peuvent être éliminées par des systèmes d'accord automatique de la cavité.

De plus, les collisions sur la paroi déplacent légèrement la fréquence de résonance. Le déplacement de fréquence dû à la présence de la paroi est de l'ordre de 10^{-11} à 300 K pour un ballon de 15 cm de diamètre [21, 22, 23]. Des études sont effectuées pour mieux comprendre la nature de cet effet [23].

Sa mesure est possible en utilisant des ballons de volume différent. La précision de la mesure est alors limitée par le manque de reproductibilité du dépôt. L'exactitude du maser à hydrogène est alors de $\pm 2 \times 10^{-12}$. Cette mesure peut être améliorée en utilisant un ballon déformable; une exactitude de l'ordre de 10^{-13} semble alors possible.

Dans le but de diminuer la fréquence moyenne des collisions sur la paroi et de se rapprocher ainsi des conditions de l'espace libre, un maser à hydrogène à très grand ballon de stockage a été développé [23]. Ce réservoir comprend deux appendices. Chacun d'eux est placé dans une cavité résonnante de la dimension habituelle. Ces deux cavités sont d'autre part couplées par l'intermédiaire d'un amplificateur à grand gain (fig. 5e). Ce dispositif est une version évoluée du résonateur à cavité à double bras et cellule de stockage.

L'exactitude des étalons de fréquence à césium et à hydrogène se situe donc autour de 10^{-12}. Il semble que cette limite puisse être portée à 10^{-13} dans l'avenir, mais au prix d'efforts importants.

Récemment, la réalisation d'un résonateur à jet d'hydrogène et cellule de stockage a été proposée [24]. La seule limitation envisagée supérieure à 10^{-14} est l'effet des collisions sur la paroi.

3.4. Comparison des Deux Techniques

Les deux étalons de fréquence qui ont actuellement les meilleures performances font appel à des techniques différentes: l'un est un discriminateur de fréquence utilisant la résonance des atomes dans le vide et l'autre un oscillateur avec cellule de stockage. Il est souhaitable qu'une comparaison directe de ces techniques soit effectuée en vue de déceler d'éventuelles erreurs systématiques, insoupçonnées jusqu'à présent, dans la mesure des fréquences de transition.

Pour cela un résonateur à jet et un maser devraient être réalisés avec le même élément. Les atomes alcalins semblent peu indiqués pour cette comparaison, essentiellement parce que leur fréquence de transition est très affectée par les collisions avec une paroi ou un gaz tampon.

Par contre, il paraît possible d'utiliser l'atome d'hydrogène pour produire un jet d'atomes très lents. Après production dans une décharge classique, les atomes pourraient être ralentis par chocs sur une paroi froide [25], ou avec un gaz à la température de l'hélium liquide. Le jet ainsi formé aurait une

vitesse correspondant à celle d'un atome de masse 100 environ, issu d'une source à la température ordinaire. Il serait ainsi possible d'obtenir, dans un appareil à jet, une résonanee suffisamment étroite pour mesurer avec une précision suffisante la fréquence de transition de l'atome d'hydrogène, et la comparer à celle que l'on obtient dans un maser à hydrogène.

4. Conclusion

Du fait de leurs performances, les étalons de fréquenee atomiques ont de nombreuses applications techniques et physiques. Dans ce dernier domaine, le maser à hydrogène joue un rôle prépondérant qui s'explique par la simplicité de l'atome d'hydrogène ou par la possibilité d'étudier les collisions avec un gaz étranger introduit dans le ballon de stockage.

A titre d'exemple, la comparaison de la valeur calculée et de la valeur mesurée de la fréquenee de transition hyperfine de l'hydrogène a conduit à une nouvelle détermination de la eonstante de structure fine [26] et l'emploi d'un maser à hydrogène en champ magnétique élevé a permis une mesure très précise du rapport des moments magnétiques de l'électron et du proton [27].

D'autre part, les progrès récents dans la stabilisation de fréquence de lasers, utilisables comme étalons de longueur, et dans la comparaison des fréquence optiques et radioélectriques laissent entrevoir la possibilité d'un étalon unique pour la fréquence, le temps et la longueur.

5. References

[1] McCoubrey, A. O., Proc. IEEE **54**, 116 (1966). Beehler, R. E., Proc. IEEE **55**, 792 (1967). McCoubrey, A. O., Proc. IEEE **55**, 805 (1967).

[2] Hellwig, H., Vessot, R., Levine, M., Zitzewitz, P., and Peters, H., Trans. IEEE Inst. (à paraitre).

[3] Larson, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23**, 1369 (1969).

[4] Mathur, B. S., Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. **158**, 14 (1967).

[5] Evenson, K. M., Wells, J. S., and Mataresse, L. M., Appl. Phys. Letters **16**, 251 (1970).

[6] Javan, A., Conference on Precision Electromagnetic Measurements Boulder (June 2–5, 1970).

[7] Ramsey, N. F., *in* Moleculars Beams, Oxford University Press (1956).

[8] Kleppner, D., Goldenberg, H. M., and Ramsey, N. F., Phys. Rev. **126**, 603 (1962).

[9] Kleppner, D., Berg, H. C., Crampton, S. B., Ramsey, N. F., Vessot, R. F. C., Peters, H. E., and Vanier, J., Phys. Rev. **138**, A972 (1965).

[10] Audoin, C., Desaintfuscien, M., Petit, P., and Schermann, J. P., Electronics Letters **5**, 292 (1969).

[11] Barger, R. L., and Hall, J. L., Phys. Rev. Letters **22**, 4 (1969).

[12] Levine, M. W., and Vessot, R. F. C., Symp. on Very Long-Baseline Interferometry, Charlottesville, Virg. (April 13–15, 1970).

[13] Lacey, R. F., Helgesson, A. L., and Holloway, J. H., Proc. IEEE **54**, 170 (1966).

[14] Rabi, I. I., Zacharias, J. R., Millman, S., and Kusch, P., Phys. Rev. **53**, 318 (1938).

[15] Holloway, J. H., and Lacey, R. F., Actes, Congrès Int. de Chronométrie, 317 (Lausanne 1964).

[16] Kartaschoff, P., and Debély, P. E., Actes, Colloque Int. de Chronométrie **A3** (Paris 1969).

[17] Holloway, J. H., and Lacey, R. F., Actes, Colloque Int. de Chronométrie **A3** bis (Paris 1969).

[18] Kastler, A., J. Phys. Rad. **11**, 255 (1950).

[19] Cérez, P., Arditi, M., and Kastler, A., C.R. Acad. Sci. **267**, B282 (1968).

[20] Goldenberg, H. M., Kleppner, D., and Ramsey, N. F., Phys. Rev. **123**, 530 (1961).

[21] Menoud, C., and Racine, J., Actes, Colloque Int. de Chronométrie A8 (Paris 1970).

[22] Vanier, J., and Vessot, R. F. C., Metrologia **6**, 52 (1970).

[23] Zitzewitz, P. W., Debély, P., and Ramsey, N. F., cette conférence.

[24] Hellwig, H., Metrologia **6**, 56 (1970).

[25] Brackman, R. T., and Fite, W. L., J. Applied Phys. **34**, 1572 (1961).

[26] Cohen, E. R., and Dumond, J. W. M., Rev. Mod. Phys. **37**, 537 (1965).

[27] Myint, T., Kleppner, D., Ramsey, N. F., and Robinson, H. G., Phys. Rev. Letters **17**, 405 (1966).

DISCUSSION

J. A. Barnes: Do you see any clearcut advantage to changing the definition of the unit of time interval, the second, to hydrogen?

C. Audoin: No, the important point is that it is an atomic definition, a decision which is good.

J. A. Barnes: You think they are comparable in accuracy and there is no reason to change?

C. Audoin: No reason.

Accurate Frequency Measurements: Survey, Significance, and Forecast

Helmut Hellwig and Donald Halford

Atomic Frequency and Time Standards Section, National Bureau of Standards, Boulder, Colo. 80302

Accurate frequency measurements in the microwave through the infrared regions as well as the imminent realization of frequency measurements in the near infrared and visible regions are surveyed, and their significance on the system of basic standards and fundamental constants is discussed. An exceedingly accurate redetermination of the speed of light, and a single primary standard for frequency, time, and length are imminent possibilities. The further possibility of one primary standard for many (if not all) of the base units exists. Traditional beam techniques, storage methods, and infrared or visible radiation molecular absorptions appear as the most promising candidates for the primary (frequency) standard.

Key words: Frequency standards; frequency/time metrology; fundamental constants; length standards; speed of light; units of measurements.

1. Introduction

The present International System of Units of Measurement (SI) is built on six base units which are related—with one exception (mass)—to fundamental properties of nature: time, length, mass, temperature, electrical current, and luminous intensity [1]. Of these the unit of time, or more exactly the unit of time interval, the second, has always been the most accurately[1] known and internationally accepted unit. The simple reason for this is that we are provided by nature with a unit of time interval: the duration of one day due to the rotation of the earth. Up to the recent past the definition of the second was based on the rotation of the earth and more recently on the revolution of the earth around the sun [2]. The accuracy of the second which can be realized by this definition approaches one part in 10^9 for extremely long observation periods (many years) [2, 3]. For shorter observation periods the accuracy is correspondingly worse. Figure 1 depicts the development of the accuracy capability of time interval standards since the advent of atomic clocks. Accuracy capability is here expressed as the one sigma combined uncertainty of all bias corrections. The bias corrections are the result of a theoretical and experimental evaluation of each particular standard, whose actual performance will always deviate to some degree from the idealized conditions which are adopted in the definition of a base unit.

Figure 1 shows that the astronomical definition held its place exclusively until 1947. In that year the first ammonia clock was put into operation by H. Lyons at the National Bureau of Standards [4]. This clock's performance did not yet surpass that of the astronomical one. However, the advent of the ammonia clock is important for two reasons: Firstly, the unit of time interval could be related for the first time to an (assumed) invariant physical constant, here to the inversion transition in the ammonia molecule, instead of being based solely on the movements of macroscopic celestial bodies which are known to have secular changes (it is possible to correct for some but not all of the secular changes). Secondly, a frequency standard was used to aid in the definition of time interval; i.e., one used the relationship

$$\tau \equiv b\nu^{-1} \tag{1}$$

and $b \equiv 1$, with τ and ν being the period and frequency of the radiation which is associated with the quantum transition. The unit of time interval, the second, can thus be defined as a certain number of periods of this radiation.[2]

The first complete atomic clock system based on a hyperfine transition in cesium was operated and evaluated in 1955 by Essen at the National Physical Laboratory [5] and exceeded the performance of previous standards by one order of magnitude. As indicated in figure 1, further improvements in the cesium atomic clock at several laboratories around the world [3], [6–15] pushed its accuracy capability to its current value of better than 10^{-12} [16–19]. This performance of cesium clocks led the 13th General Conference on Weights and Measures in 1967 to accept the following definition: "The second is the duration of 9192 631 770 periods of the radiation corresponding to the transition between the two

[1] By accuracy, we mean the degree to which a physical measurement or its measuring device conforms to a specified definition.

[2] There could be separate standards for time interval and for frequency. The constant b in eq (1) would then be a fundamental constant in a sense quite similar to the speed of light, and b would have the dimension Hz·s.

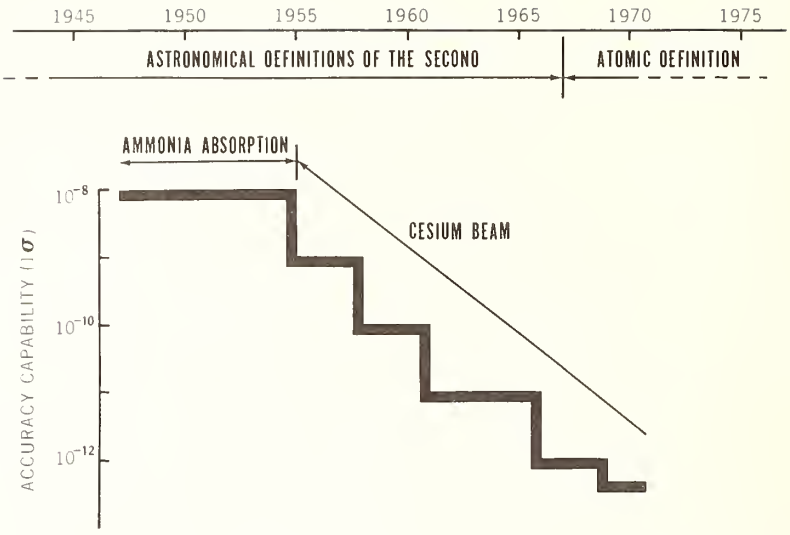


FIGURE 1. *Historical development of the accuracy capability of the leading quantum electronic time interval/frequency standards, up to Summer 1970.*

hyperfine levels of the ground state of the cesium-133 atom." The unit of frequency, the hertz, is then defined by eq (1) with $b \equiv 1$ Hz·s.

2. Accurate Frequency Measurements: Principles and Methods

The basic standard for time interval, the cesium atomic beam apparatus, is not only the most accurate frequency source, but also the most accurate of all basic standards by a considerable margin as is illustrated in figure 2. This graph depicts the accuracy capability of all six standards of the SI Base Units [1, 20]. The values for time, length, current, temperature, and luminous intensity describe the ability of actual instruments to realize the definition of the corresponding base unit. They were obtained from theoretical and experimental evaluations of bias corrections and from national and international intercomparisons. The only base unit which is still defined by an artifact is the kilogram. Here the accuracy capability is the ability to compare masses with this artifact and to preserve it unchanged. It must be noted that the precision[3] of relative measurements in any of these base units may be considerably better than the quoted accuracy capability of the corresponding basic standard. In addition to being the most accurate kind, frequency/time metrology is the most precise of the many kinds of metrology (e.g., length, mass, force, pressure, resistance, current). Accurate and precise frequency measurements can be easily instrumented and can be highly automated. The versatility of frequency measurement techniques has led to their wide usage in metrology in general. Radio broadcasts of accurate frequency and time signals are available worldwide.

It is obvious from figure 2 that measurements which are based on time interval or frequency determinations have the potential of exceeding by far the accuracy of any measurement involving the five other base units. Some physical constants can therefore be measured with extreme accuracy by using time interval and frequency methods. In particular this is true for the measurement of quantum transitions in atoms and molecules.

The availability of an exceedingly accurate standard is, however, only one prerequisite for an accurate measurement. In order to utilize the high accuracy capability of a time interval/frequency standard, the system which is to be measured has to be brought under such experimental control that its measurement yields an accuracy which approaches that of the standard. In the limit, the system under study will have to show properties which are charac-

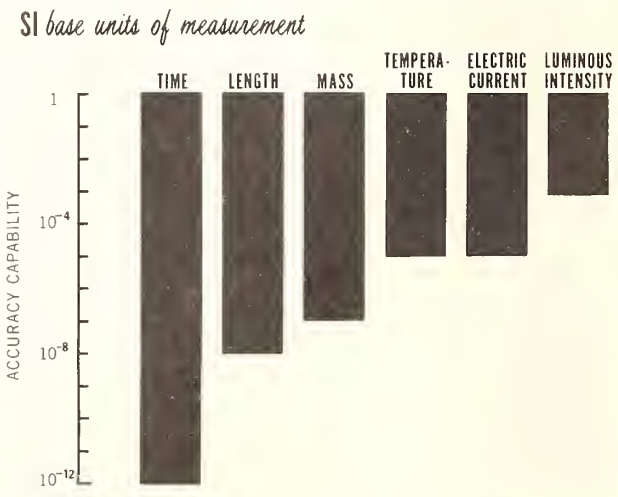


FIGURE 2. *Comparison of the current accuracy capability of the six basic standards of the International System of Units (SI).*

[3] By precision, we mean the reproducibility of a specified set of measurements taken as a time series.

teristics of a time interval/frequency standard itself. Therefore, the very systems which yield the most accurate measurements are simultaneously candidates for new time interval/frequency standards. We will discuss this aspect later in this paper.

A system intended for the measurement of a transition frequency of an atom or molecule involves several steps of physical and technical processing which lead to bias corrections and corresponding uncertainties. These steps can be classified into three groups: (1) particle preparation, (2) particle confinement, and (3) particle interrogation [21].

There is no fundamental necessity to prepare the particles. However, often it is desirable to select only certain energy states or to achieve a desired population distribution of the energy levels. This can be done by spatial state selection, optical pumping, etc. Care must be taken to minimize perturbing effects.

A fundamental aim of precise measurements is an observation time as large as possible. Particle confinement is used to achieve coherent interaction times between the particles and the interrogating radiation which are as long as possible without introducing undue perturbations or excessive loss of signal. The confinement technique also is of considerable importance in the reduction of the Doppler effect, which is the most severe limitation in simple gas cell absorption measurements. Various confinement techniques such as a storage vessel with coated walls, ion storage, the storage in a cell filled with buffer gases, or the traveling particle beam are possible.

The particles have to be interrogated. Usually a resonant structure (e.g., cavity or interferometer) is used to enhance the interaction and to provide a spatially well-defined interaction region. The interrogation process itself may introduce perturbations such as the Bloch-Siegert effect [22] and photon recoil [21, 23].

We will now discuss in some detail the three methods which stand out as leading to the most accurate measurements. It is not surprising that all three have in common a significant reduction of the Doppler effect limitation.

2.1. Particle Beams

In a particle beam apparatus, particles emerge from a source, forming a beam which travels through vacuum as indicated in figure 3. A polarizer (spatial state selector, optical pump, . . .) creates a certain, more desirable population distribution of the energy levels. The beam then enters the region of interrogation and is "confined" there for the duration of the transit time. The interrogation by radiation of suitable intensity and a suitable frequency which is compared to the frequency standard results in a change of the population distribution and a change of the intensity of the transmitted or reflected radiation. Two modes of detection are therefore possible. One is the analysis of the population distribution by counting the number of particles in a selected energy state. An analyzer, similar to the polarizer, and a particle detector are then necessary. An example of this mode of detection is shown in figure 3 [24]. The other mode of detection, which is not shown in figure 3, involves the monitoring of the radiation intensity [25, 26]. Under certain conditions, particularly at sufficiently high beam intensity, the system converts to a frequency generator (maser, laser oscillator) [25]. The output frequency is then detected and measured by comparison with a frequency standard.

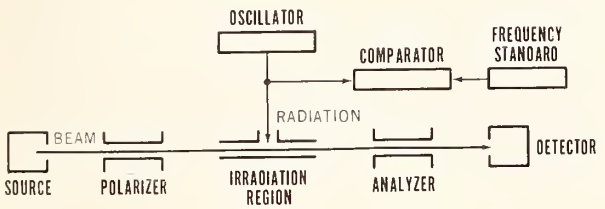


FIGURE 3. *Accurate frequency measurements: particle beam method*

In the particle beam the Doppler effect is greatly reduced by the narrow, unidirectional beam. If care is taken in the design of the resonance structure so that the beam does not encounter net radiation power traveling parallel to the beam direction, then limitations due to the first-order Doppler effect can be virtually eliminated [24]. In order to correct for the second-order Doppler effect, the particle speed must be adequately known. Particle preparation does not introduce frequency bias if it is spatially separated from the interrogation region.

2.2. Storage of Particles

In a storage device, particles are stored in a vessel which is located within the resonance structure as shown in figure 4. Particle preparation may be done simultaneously with the interrogation, e.g., by optical pumping and monitoring the intensity of the transmitted pump radiation as indicated in figure 4 [27]. However, this introduces frequency shifts which are difficult to evaluate [28]. It is therefore advantageous to separate the preparation and interrogation regions and to use the storage principle in connection with the polarizing and analyzing technique of figure 3. In other words, the storage

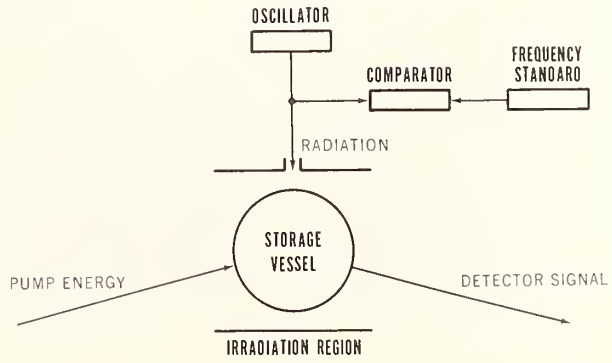


FIGURE 4. *Accurate frequency measurements: particle storage method.*

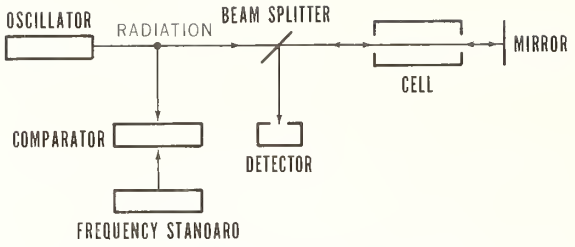


FIGURE 5. *Accurate frequency measurements: saturated absorption method*

device of figure 4 is inserted between polarizer and analyzer of the system in figure 3 thus forming a storage beam device [29, 30]. Again we have the option to adjust the system parameters (beam intensity, pumping strength, ...) such that self-oscillations are possible [28, 31].

The advantage of the storage technique is the increase in the confinement time leading to a very sharp line. The first-order Doppler effect also can be virtually eliminated if the movement of the particles is confined to a region of less than half the wavelength of the interrogating radiation. This is easy for the case of microwave frequencies but is more difficult at very short wavelengths. However, buffer gases can sufficiently restrict the particle movements. Also, buffer gases [28] and coating of the walls of the storage vessel [31] are used to reduce frequency shifts and relaxation processes due to wall collisions. The second-order Doppler shift can be calculated to a high degree of accuracy from the temperature of the storage vessel since the kinetic energy of the stored particles will be in thermal equilibrium with the storage vessel.

2.3. Saturated Absorption

As shown in figure 5, in a saturated absorption device, the radiation of an oscillator is passed in opposite directions through a cell which is filled with the particles under study. The transmitted radiation intensity is monitored. Those particles with a near-zero velocity component parallel to the radiation propagation vector experience a nonlinear, enhanced interaction with the radiation field ("Lamb dip") [32, 33]. The system parameters can be adjusted so that those particles which have a significant velocity component parallel to the radiation beam are not interrogated. This reduces considerably the first-order Doppler effect and also excludes from the interrogation most of the particles which suffered a collision with other particles.[4] The line width is thus ultimately given by the transit time of particles across the radiation beam. The second-order Doppler correction can be obtained from the gas cell temperature. At higher frequencies the saturated absorption method may be limited by photon recoil effects which cause the emitted frequency ν_E to be different from the absorbed frequency ν_A by the fractional amount [21, 23]

$$(\nu_A - \nu_E)/\nu = h\nu/mc^2, \qquad (2)$$

where ν is the average of ν_A and ν_E, m is the mass of the particle, h is Planck's constant, and c is the speed of light.

3. Survey of Accurate Frequency Measurements

For figure 6 we have selected the most significant of the many published accurate frequency measurements and have plotted the published one sigma accuracy versus the location of the transition in the electromagnetic spectrum. The dashed bars in the cases of CH_4 and I_2 indicate that their frequencies have not yet been measured.[5] This is because frequency multiplication has not yet succeeded in reaching to these frequencies above 30 THz. However, advances by several groups, in particular by K. Evenson of NBS [34] and A. Javan of MIT [35], indicate that success is imminent. The values indicated for the accuracy of the anticipated measurements of CH_4 and I_2 are based on the evaluation of the beat frequency between two lasers which were independently frequency-locked to the CH_4 transition [32] and on similar studies together with absolute wavelength measurements in the case of I_2 [33]. Table 1 contains information supplementary to figure 6.

The measurements of HCN, H_2O, and CO_2 were made by frequency multiplication in a metal-metal point contact diode. The comparatively low accuracy of these values is due mainly to the fact that the molecular transition was investigated in a laser oscillator. This technique introduces large uncertainties because of Doppler effects, pressure effects, resonator frequency pulling, etc. Application of the saturated absorption method would considerably increase the measurement accuracy as is demonstrated by the CH_4 and I_2 experiments.

The storage of ions is a powerful and very promising spectroscopic tool as is demonstrated by the $^3He^+$ measurement. However, several experimental parameters are not yet fully understood, and they need further investigation, especially the kinetic energy of the stored ions (Doppler effect) [36].

The methods on which the most accurate measurements are based were already discussed. They are the saturated absorption (CH_4, I_2), the storage vessel (H, D, T), and particle beams (Tl, NH_3, H_2S, Cs). The cesium particle beam tube serves as the primary standard for the base unit of time interval and frequency. It has been the most accurate measuring device for the past fifteen years (fig. 1). Its present accuracy capability, shown as the horizontal line across the bottom of figure 6, is five parts in 10^{13} [15, 17, 18, 19].

[4] Particle collisions lead to changes in particle speed and direction. A strong collision removes the particle from the interrogation process because the collision introduces, with high probability, a significant velocity component parallel to the radiation beam.

[5] To date, their frequencies have been determined to about six digits by length metrology, but not at all by frequency metrology.

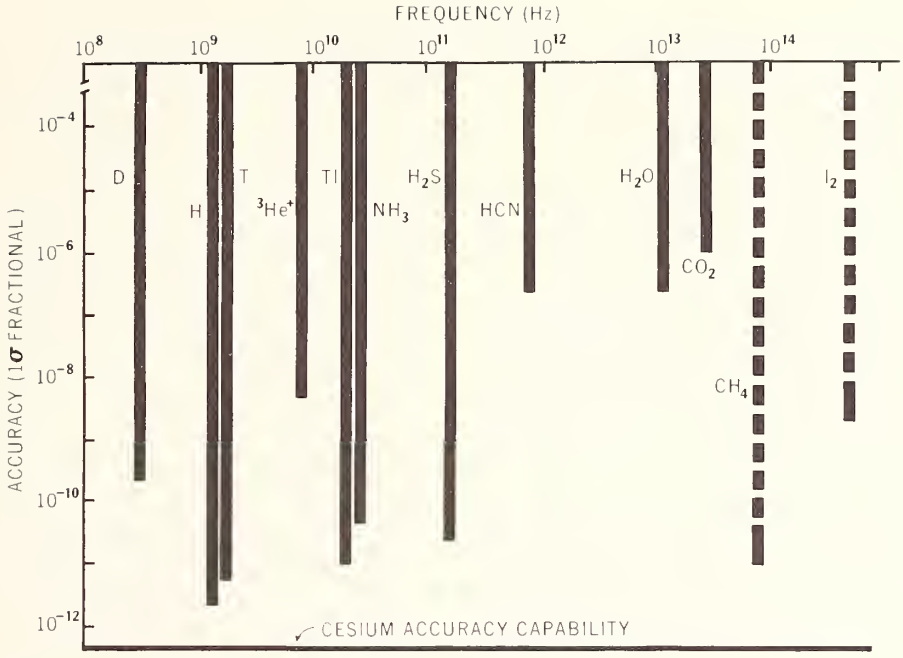


FIGURE 6. *A survey of selected accurate frequency measurements throughout the electromagnetic spectrum.*
The accuracies reflect the state-of-the-art as of Summer 1970. For CH_4 and I_2 the anticipated values are indicated.

4. Significance and Impact of Accurate Frequency Measurements

We can identify four general areas where accurate frequency measurements are significant and where some impact on future scientific and technological developments is foreseen. These are summarized in table 2. Within the scope of this paper only the last area, fundamental constants and basic standards, is of importance; we will discuss it in detail and only briefly explain the first three items.

Metrology and applications include radar ranging especially over long (planetary) distances; planetary exploration in general; earth and space navigation, where timekeeping over weeks, months, or even years without resynchronization is vital; telecommunication aspects including high bit rates and better usage of the electromagnetic spectrum; and aircraft collision avoidance systems which can be based perhaps with advantage on time domain techniques which employ accurate clocks.

Accurate frequency measurements throughout the infrared and visible region will greatly increase our knowledge of the structure of atoms and molecules. Spectroscopic constants such as transition frequencies, g-factors, Stark coefficients, rotational distortions, etc., will be accessible to measurements of unprecedented precision and accuracy.

TABLE 1. *Survey of accurate frequency measurements*

Particle	Transition	Technique	Frequency[a]	Accuracy	References
D	F, $m_F=\frac{3}{2}, \frac{1}{2}\leftrightarrow\frac{1}{2}, -\frac{1}{2}$	Storage maser	327 384 352.51 Hz	2×10^{-10}	[37]
H	F, $m_F=1$, $0\leftrightarrow0$, 0	Storage maser	1 420 405 751.768 Hz	2×10^{-12}	[38, 39]
T	F, $m_F=1$, $0\leftrightarrow0$, 0	Storage maser	1 516 701 470.809 Hz	5×10^{-12}	[40]
$^3He^+$	F, $m_F=1$, $0\leftrightarrow0$, 0	Ion storage	8 665 649 905 Hz	6×10^{-9}	[41]
^{205}Tl	F, $m_F=1$, $0\leftrightarrow0$, 0	Particle beam	21 310 833 945.9 Hz	1×10^{-11}	[42, 43]
$^{15}NH_3$	J, K=3, 3 inversion	Beam maser	22 789 421 731 Hz	5×10^{-11}	[44]
$H_2{}^{32}S$	$1_{-1}\leftrightarrow1_1$	Particle beam	168 762 762 373 Hz	2×10^{-10}	[45]
$H^{12}C^{14}N$	$110\leftrightarrow040$	Laser	0.890 7606 THz	2×10^{-7}	[46, 47]
$H_2{}^{16}O$	$001\leftrightarrow020$	Laser	10.718 073 THz	2×10^{-7}	[47]
$^{12}C^{16}O_2$	P(18), $001\leftrightarrow100$	Laser	28.359 800 THz	1×10^{-6}	[34]
$^{12}C^{16}O_2$	R(12), $001\leftrightarrow020$	Laser	32.176 085 THz	6×10^{-7}	[50]
$^{12}CH_4$	P(7), ν_3 band	Saturated absorption	≈88 THz	(1×10^{-11})	[32]
I_2	R(127), 11–5 band of electr. trans. $B\,{}^3\Pi^+{}_{0_u}\leftrightarrow X\,{}^1\Sigma^+{}_{0_u}$	Saturated absorption	≈474 THz	(2×10^{-9})	[33]

[a] By definition, $\nu_{Cs}\equiv9192\,631\,770$ Hz.
[b] The frequencies of several other transitions in water vapor lasers have been measured such as the 2.5 THz (118 μm) line in H_2O [48] and the 3.6 THz (84 μm) line in D_2O [49].

TABLE 2. *Areas of Impact*

Metrology and Applications to Technology
Spectroscopic Constants
General Theory of Relativity
Fundamental Constants and Basic Standards

Tests of the general theory of relativity involving differences of coordinate time and frequency at locations of different gravitational potential will be made with clocks placed in satellites or on other celestial bodies. An improvement in accuracy of about one order of magnitude over our present accuracy capability will permit some tests to be conducted on the Earth's surface.

The ability to perform accurate frequency measurements throughout the electromagnetic spectrum will have a considerable impact on the system of fundamental constants and basic standards. Figure 6 illustrates that frequency measurements have already been made in the terahertz region and that the ability to compare the standard of time interval and frequency (currently ^{133}Cs) with the standard of length (currently ^{86}Kr) will soon be a reality. The relationship between these standards involves the speed of light c,

$$\lambda = c\nu^{-1}, \tag{3}$$

where λ is the vacuum wavelength and ν is the frequency of the radiation. Equation (3) as applied to the direct comparison of length and frequency (time interval) will lead to a determination of the speed of light with unprecedented accuracy [34], which will be limited only by the ability to perform an interferometric length measurement (compare fig. 2).

Equation (3) is not the only "simple" relationship between basic standards involving a fundamental constant. The Josephson phenomenon [51] provides us with a relationship between the electric potential difference V (relating to the standard of electrical current) across a superconducting weak link and the corresponding Josephson oscillation frequency ν

$$V = (h/2e)\nu. \tag{4}$$

Equation (4) has already served to determine Planck's constant h divided by the electronic charge e with unprecedented accuracy [52, 53] which led to a refitting of the whole system of fundamental constants [54]. We can imagine further simple relationships between basic standards and the time interval/frequency standard involving fundamental constants, such as a relationship for the mass m

$$m = f(\nu) \tag{5}$$

or for the temperature T

$$T = g(\nu). \tag{6}$$

Accurate frequency measurements based on eqs (3) to (6) could therefore lead to a more accurate knowledge of the fundamental constants. This increased accuracy in turn would allow a more sensitive search for possible spatial and secular variations of the fundamental constants [55].

It is a consequence of historical development and experimental expertise that we have a set of base units and corresponding standards such as the SI; no fundamental physical principle is involved in this choice. It is already possible to compare spectroscopic data in the infrared and visible region using frequency measurement techniques with a precision far exceeding that of interferometric (length measurement) techniques.

In view of the imminent success of frequency multiplication from the microwave region into the infrared and visible regions it seems to be in order to question the need for a separate standard of length. Length measurements could be related to the base unit for time interval and frequency via eq (3). The speed of light would then be a defined constant[6] [56, 57, 58]. The defined value of c would be chosen to be compatible with its previous best experimental value.

In the limit this philosophy would lead to a base set of fundamental constants with defined values. The most accurate standard, which is today the cesium beam frequency standard, would then serve as *The Standard*, and all other units of measurement could be derived using relationships such as eqs (3) to (6) with c, $h/(2e)$, and some yet undetermined functions like $f(\nu)$ and $g(\nu)$ as defined entities. These possibilities add excitement to the work which is being done on frequency standards.

5. Summary

In the preceding sections, we pointed out that frequency/time metrology is currently the most precise and accurate kind of metrology, and we outlined the significance of accurate frequency measurements and their impact on technology and physics, and in particular on the system of fundamental constants and basic standards. Possibly being *The Standard* of the future, the primary frequency standard should get special attention. At present we see several ways of pushing its accuracy capability beyond the present one sigma performance of 5×10^{-13} of the laboratory cesium beam tube [21].

We discussed the three most promising methods which allow the most accurate frequency measurements: particle beam, particle storage, and saturated absorption. We also mentioned that naturally these very same methods are the most promising ones for the development of future standards of even better accuracy.

In table 3, experimental realizations of these methods are listed together with evaluated accuracy capabilities and projected performances. Each pro-

[6] Compare with footnote 2.

Table 3. *Promising candidates for primary frequency standards*

Method	Device	ν(Hz)	Accuracy Capability (one sigma)	
			Present performance	Projected performance
Particle beam	Cesium beam	9.2×10^{9}	5×10^{-13} [19]	1×10^{-13}
	Iodine beam	5.8×10^{14}	— [26]	better than 10^{-12}
Particle storage	Hydrogen maser	1.4×10^{9}	2×10^{-12} [38]	better than 10^{-12}
	Hydrogen storage beam	1.4×10^{9}	— [30]	better than 10^{-12}
	$^{3}He^{+}$ ion storage	8.7×10^{9}	6×10^{-9} [27]	better than 10^{-11}
Saturated absorption	Methane cell	8.8×10^{13}	1×10^{-11} [32]	better than 10^{-12}
	Iodine cell	4.7×10^{14}	2×10^{-9} [33]	better than 10^{-11}

jected performance is listed as better than some quoted value. This value is believed to represent a minimum performance which almost certainly can be surpassed based on our present experimental and theoretical knowledge; however, we can not predict with sureness by how far. In this sense, a technique quoted as "better than 10^{-11}" could well surpass one quoted as "better than 10^{-12}." For a more detailed discussion we recommend the cited literature and we suggest reference [21] for a critical comparison and prognostic.

Perhaps a uniquely superior method of accurate frequency measurement will be singled out by additional experimental and theoretical work, or perhaps these three methods will each improve significantly so as to remain mutually competitive. Very likely, in the continuing progress of knowledge and technical abilities, new possibilities will continue to appear.

6. References

[1] Terrien, J., Metrologia **1**, 15 (1965).
[2] Kovalevsky, J., Metrologia **1**, 169 (1965).
[3] Markowitz, W., Hall, R. G., Essen, L., and Parry, J. V. L., Phys. Rev. Letters **1**, 105 (1958).
[4] Lyons, H., Annals of the N. Y. Acad. of Sci. **55**, 831 (1952).
[5] Essen, L., and Parry, J., Nature **176**, 280 (1955).
[6] Kalra, S., Bailey, R., and Daams, H., Can. J. Phys. **36**, 1442 (1958).
[7] Mainberger, W., Electronics **31**, 80 (1958).
[8] McCoubrey, A., Proc. 12th Annual Symp. on Freq. Control, Ft. Monmouth, N. J., 648 (1958).
[9] Kartaschoff, P., Bonanomi, J., and De Prins, J., Helv. Phys. Acta **33**, 969 (1960).
[10] Kartaschoff, P., IRE Trans. on Instr. **I-11**, 224 (1962).
[11] Beehler, R. E., and Glaze, D. J., IEEE Trans. on Instr. and Meas. **IM-15**, 48 (1966).
[12] Mungall, A. G., Bailey, R., and Daams, H., Metrologia **2**, 97 (1966).
[13] Bodily, L., Hartke, D., and Hyatt, R., Hewlett-Packard J. **17**, 13 (1966).
[14] Vessot, R., Peters, H., Vanier, J., Beehler, R., Halford, D., Harrach, R., Allan, D., Glaze, D., Snider, C., Barnes, J., Cutler, L., and Bodily, L., IEEE Trans. on Instr. and Meas. **IM-15**, 165 (1966).
[15] Mungall, A. G., Bailey, R., Daams, H., and Morris, D., Metrologia **4**, 165 (1968).
[16] Kartaschoff, P., and Debély, P. E., Proc. of the Colloque International de Chronométrie, Paris, p. A3-1 (1969).
[17] Becker, G., Fischer, B., Kramer, G., and Müller, E. K., Proc. of the Colloque International de Chronométrie, Paris, p. A2-1 (1969).
[18] Becker, G., Fischer, B., Kramer, G., and Müller, E. K., PTB-Mitteilungen **2**, 77 (1969).
[19] Glaze, D. J., Proc. of the XVI General Assembly of URSI, Ottawa, Canada, August 1969, entitled Progress in Radio Science, 1966–1969, to be published; also published in IEEE Trans. on Instrumentation and Measurement **IM-19**, 156 (1970).
[20] Terrien, J., Metrologia **2**, 55 (1966).
[21] Hellwig, H., Metrologia **6**, 118 (1970).
[22] Shirley, J. H., J. Appl. Phys. **34**, 783 (1963).
[23] Kolchenko, A. P., Rautian, S. G., and Sokolovskii, R. I., Zh. Eksp. Teor. Fiz. **55**, 1864 (1968) (English translation in Soviet Phys. JETP **28**, 986 (1969).
[24] Ramsey, N. F., Molecular Beams. Oxford (1956).
[25] Shimoda, K., Wang, T. C., and Townes, C. H., Phys. Rev. **102**, 1308 (1956).
[26] Ezekiel, S., and Weiss, R., Phys. Rev. Letters **20**, 91 (1968).
[27] Dehmelt, H. G., Phys. Rev. **103**, 1125 (1956).
[28] Davidovitz, P., and Novick, R., Proc. IEEE **54**, 155 (1966).
[29] Kleppner, D., Ramsey, N. F., and Fjelstad, P., Phys. Rev. Letters **1**, 232 (1958).
[30] Hellwig, H., Metrologia **6**, 56 (1970).
[31] Kleppner, D., Berg, H. C., Crampton, S. B., Ramsey, N. F., Vessot, R. F. C., Peters, H. E., and Vanier, J., Phys. Rev. **138**, A972 (1965).
[32] Barger, R. L., and Hall, J. L., Phys. Rev. Letters **22**, 4 (1969).
[33] Hanes, G. R., and Baird, K. M., Metrologia **5**, 32 (1969).
[34] Evenson, K. M., Wells, J. S., and Matarrese, L. M., Appl. Phys. Letters **16**, 251 (1970).
[35] Daneu, V., Sokoloff, D., Sanchez, A., and Javan, A., Appl. Phys. Letters **15**, 398 (1969).
[36] Church, D., and Dehmelt, H. G., J. Appl. Phys. **40**, 3421 (1969).
[37] Larson, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23**, 1369 (1969).
[38] Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. Letters **11**, 338 (1963).
[39] Hellwig, H., Vessot, R. F. C., Levine, M. W., Zitzewitz, P. W., Allan, D. W., and Glaze, D. J., IEEE Trans. on Instr. and Meas. **IM-19**, 200 (1970).
[40] Mathur, B. S., Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. **158**, 14 (1967).
[41] Fortson, E. N., Major, F. G., and Dehmelt, H. G., Phys. Rev. Letters **16**, 221 (1966).
[42] Bonanomi, J., IRE Trans. on Instr. **I-11**, 212 (1962).
[43] Beehler, R. E., and Glaze, D. J., IEEE Trans. on Instr. and Meas. **IM-15**, 55 (1966).

[44] De Prins, J., Menoud, C., and Kartaschoff, P., Rev. Sci. Instr. **32**, 1267 (1961).
[45] Cupp, R. E., Kempf, R. A., and Gallagher, J. J., Phys. Rev. **171**, 61 (1968).
[46] Hocker, L. O., and Javan, A., Phys. Letters **25A**, 489 (1967).
[47] Evenson, K. M., Wells, J. S., Matarrese, L. M., and Elwell, L. B., Appl. Phys. Letters **16**, 159 (1970).
[48] Frenkel, L., Sullivan, T., Pollack, M. A., and Bridges, T. J., Appl. Phys. Letters **11**, 344 (1967).
[49] Hocker, L. O., Small, J. G., and Javan, A., Phys. Letters **29A**, 321 (1969).
[50] Sokoloff, D., and Javan, A., Bull. Am. Phys. Soc. Series II **15**, 505 (1970).
[51] Josephson, B. D., Adv. Phys. **14**, 419 (1965).
[52] Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177**, 639 (1969).
[53] Petley, B. W., and Morris, K., Metrologia **6**, 46 (1970).
[54] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[55] Dirac, P. A. M., Proc. Royal Soc. London **A165**, 199 (1938).
[56] Simkin, G. S., Izmeritel'naya Tekhnika **No. 10**, 6 (1968) (English translation in Measurement Techniques **No. 10**, 1308 (1968)).
[57] Townes, C. H., Science **149**, 831 (1965).
[58] Essen L., Nature **180**, 137 (1957).

DISCUSSION

K. M. Baird: There were a number of things, especially on your slide indicating accuracy of various frequency standards: for the two in the infrared and the visible you gave then, the CH_4 and the iodine. I wonder why you stopped the dotted lines where you did.

You had, for example, 10^{-9} for iodine. Does this refer to a comparison with the microwave standards? And if so, it's the process of comparison rather than the accuracy of the standard and so perhaps you're a little unfair to those. If you're going to dot them anyway, they should go down through the bottom of the screen maybe.

H. Hellwig: I originally hesitated to put them in; therefore, the dotted lines. What the dotted lines mean is what we believe can at least be achieved if multiplication is possible. Let me state that very clearly. It can go, as you say, very far to the bottom. But right now we are unable to evaluate them further down. The value I stated there was taken from your paper where you are quoting reproducibility.

K. M. Baird: I certainly agree they should be dotted, because it is very early to guess the limiting absolute accuracy.

H. Hellwig: Yes.

K. M. Baird: But if you're going to dot them anyway, I thought they might go down farther. (*Laughter.*)

H. Hellwig: I think I took care of that by the last table.

K. M. Baird: Yes. Indeed, yes. There was another thing I would like to comment on. If we are working on standards in the sense of getting uniformity, there is always the question of how to express accuracy because your accuracy increases with a decreasing number.

H. Hellwig: Yes.

K. M. Baird: And it's really error you're indicating.

H. Hellwig: It's just a question of representation.

K. M. Baird: Yes.

H. Hellwig: We thought about that in preparing these slides, and we thought first of, say, putting in the number of digits to which it could be determined, and then we thought we should stick with the ten-to-the-minus figure and do credit to the mathematical aspect by plotting them downwards.

K. M. Baird: I have it all mixed up in my paper. (*Laughter.*)

A. H. Cook: May I make two supplementary points to a very interesting paper. I think it's well worthwhile saying to people who are not geodesists or involved in celestial mechanics that one already uses the velocity of light as a standard of length. All our knowledge of the size of the Earth and the distances of heavenly bodies is based on the velocity of light and not on a standard of length. Secondly, it seems to me that when one is dealing with those technological matters which Dr. Branscomb mentioned earlier, the vastly important advantage of using frequency as one's fundamental standard is that one can broadcast it over a radio station and you don't have to carry it around individually with you.

H. Hellwig: Thank you. May I cite my coauthor? He likes to say that we are already broadcasting the standard volt.

A. H. Cook: Yes.

J. S. Thomsen: I just want to clarify the distinction that you made earlier in your talk between time and frequency which you said has arisen in recent years. Does this simply imply that you can compare two frequencies more accurately than you can measure any interval in terms of number of cycles?

H. Hellwig: Yes. Suppose you have a time interval measurement device, an actual one. The question there would be how precisely can we locate events on a time scale. And I think the best which can be done is now in the pico-second range, whereas with frequency we can do 10^{-15}. So we have a few orders of magnitude difference. Time measurement devices are kind of inconvenient if you are driving for high accuracy.

J. S. Thomsen: If you're talking about long times, then can't you get time interval to the same accuracy that you know frequency?

H. Hellwig: I think so, yes.

D. Halford: Very long time.

H. Hellwig: Very long time.

J. S. Thomsen: Very long time.

H. Hellwig: But most measurements should not wait that long.

F. G. Major: In the interests of clarifying the status of the ion storage technique in relation to the

others, I think I should point out, as many of you probably already know, that the helium ion is cited simply because, of course, it has been done, but it was done for entirely different reasons than making a frequency standard. It was done because it was a hydrogenic system, and it was done in a way to demonstrate the feasibility of that technique. Hopefully, as this technique is developed, there will be forms of ion storage standards which far exceed any numbers indicated on that chart.

H. Hellwig: Yes, I fully agree with Dr. Major. And let me say again or supplement what he said: So far there are only publications on ion storage from the group around Dehmelt at Washington to which Dr. Major once contributed also. And at this group there was no intention to use this device as a frequency standard.

So I agree fully, if there is more emphasis on this aspect we can expect, I think, what might be spectacular progress.

Recent Developments Affecting the Hydrogen Maser as a Frequency Standard *

R. F. C. Vessot and M. W. Levine

Smithsonian Astrophysical Observatory and Harvard College Observatory, Cambridge, Mass. 02138

and

P. W. Zitzewitz, P. Debely,*** and N. F. Ramsey**

Lyman Physics Laboratory, Harvard University, Cambridge, Mass. 02138

The processes limiting the accuracy and stability of hydrogen-maser frequency-standard systems are described. Recent developments of the maser systems in the following areas are discussed:

(1) The temperature dependence of wall shifts.
(2) The use of the zero wall-shift temperature to eliminate the wall shift.
(3) The deformable storage bulb technique for varying the wall-collision rate.
(4) The large-bulb maser.
(5) Electronic phase-lock systems and cavity-tuning servos.

Key words: Accurate frequency standards; accurate time standards; frequency stability; maser oscillators; oscillator phase locks; quantum electronics.

1. Introduction

1.1. Description of the Maser

The atomic hydrogen maser was invented in 1960 by Goldenberg, Kleppner, and Ramsey [1]. Since this time, the maser has been under continual development as a frequency standard. At present, when used to control a crystal oscillator with a phase-lock system, it has been found to exceed that of all other known systems for sampling periods ranging up to at least 10^4 s.

Figure 1a shows a schematic diagram of the maser; figure 1b, the energy levels of atomic hydrogen; and figure 1c, a cutaway drawing of a compact maser oscillator packaged in a cylindrical form 20 inch in diameter and 32 inch in length. The hydrogen maser is the first successful maser to use a confining technique to keep a constantly replenished group of state-selected atoms in a region of rf field. The stored atoms, interacting with their rf magnetic field at their resonance frequency, are stimulated to emit continuous radiation. When this radiation is strong enough to sustain a coherent rf magnetic field at a high enough level, the system becomes self-oscillating. In practice, the atoms are stored in a bottle located so that it occupies a volume of rf magnetic field of nearly constant phase and direction in a resonant cavity.

Limitations in the stability and accuracy of the maser arise from several causes. The fundamental stability limitations are imposed by the oscillator line width, thermal noise within the line width, and thermal noise within the bandwidth of the equipment receiving the maser signal. Further stability limitations arise from systematic effects that perturb the oscillator frequency.

Accuracy limitations arise from incomplete knowledge of the magnitude of four contributions to frequency shift: (1) the magnetic field offset $\Delta\nu_m$, (2) the second-order Doppler shift $\Delta\nu_T$, (3) the effect of cavity mistunings $\Delta\nu_s$, and (4) the effect of atomic-hydrogen collisions with the walls of the storage bottle (wall shift) $\Delta\nu_w$. These shifts, when added to the unperturbed hydrogen hyperfine frequency separation ν_0, give $\nu_{\text{out}}=\nu_0+\Delta\nu_m+\Delta\nu_T+\Delta\nu_s+\Delta\nu_w$, the maser output frequency.

At present, the most uncertain correction is the wall-shift effect.

1.2. Stability Limitations

Stability is restricted by thermal noise. The expression giving the effect of noise-frequency components within the line width of the oscillator and in the bandwidth of the receiver is given by Cutler and Searle [2]:

$$\sigma(\tau)=\left[\frac{kT}{2P}\left(\frac{F\omega_B Q_e}{\omega_0^2\tau^2 Q_{e,l}}+\frac{1}{Q_l^2\tau}\right)\right]^{1/2}. \qquad (1)$$

* This work was supported in part by the National Aeronautics and Space Administration, the National Bureau of Standards, the National Science Foundation, and the Office of Naval Research.

** Currently at the University of Western Ontario, London, Ontario.

*** Currently at the Laboratoires Suisses de Recherches Horlogères, Neuchatel, Switzerland.

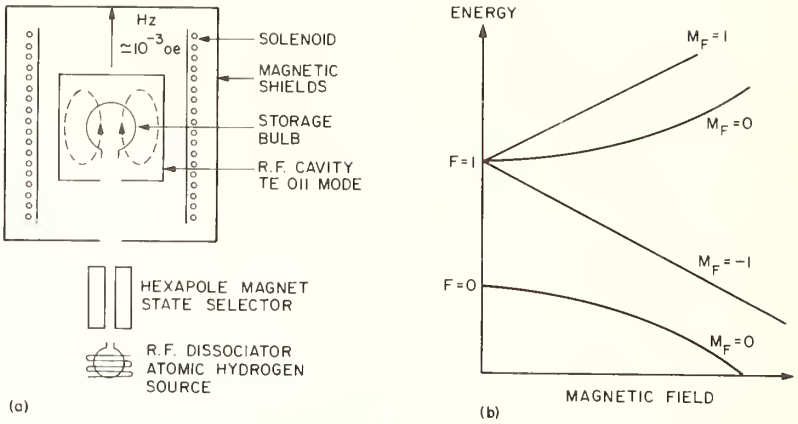


FIGURE 1. *(a) Schematic diagram of the hydrogen maser; (b) energy-level diagram of atomic hydrogen.*

Here, $\sigma(\tau)$ is the fractional rms frequency deviation averaged over a time interval τ; kT is the thermal noise power per unit bandwidth; ω_B is the receiver half-bandwidth (single-tuned band-pass characteristic); P is the power delivered by the atoms to the cavity; Q_l is the atomic line Q given by $Q_l = \omega_0/2\gamma_2\tau$; $Q_{c,l}$ is the loaded cavity Q; and Q_e is the external cavity Q.

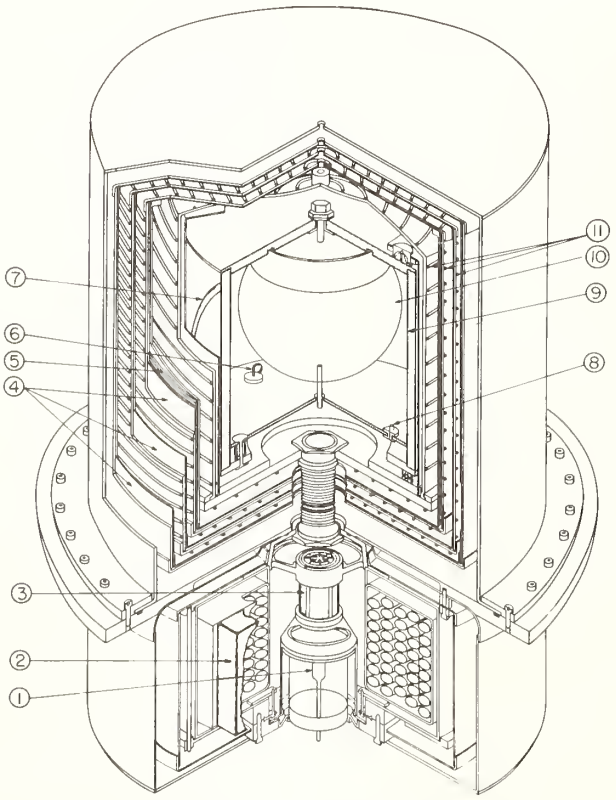


FIGURE 1c. *Hydrogen maser.*
Item 1, rf discharge tube; item 2, ion pump; item 3, hexapole magnet; item 4, magnetic shield; item 5, solenoid; item 6, tuning loop; item 7, low-frequency transition coil; item 8, output coupling loop; item 9, cavity structure; item 10, quartz storage bulb; item 11, oven heaters.

Generally, for $\tau < 30$ s, the first term in eq (1) dominates and the stability function $\sigma(\tau)$ goes as τ^{-1}. At longer periods, the effect due to the second term, resulting from thermal noise within the oscillator line, should dominate. The $\tau^{-1/2}$ behavior thus predicted has not yet been observed because it has been obscured by systematic effects that have leveled off $\sigma(\tau)$ at about 7×10^{-15}. Figure 2 shows the two-sample variance [3] as a function of averaging time τ for the maser. The lower curve shows the theoretical behavior (as described in eq 1) that applies to the maser and to the system used to obtain the data.

The single most important aspect of the storage-bulb maser technique is the ability to store an aggregate of atoms with an average velocity very close to zero in the standing wave of a cavity so that they can interact for an extended period of time, $\Delta\tau$, with the in-phase components of the rf magnetic field in the cavity. The narrow resonance line width $\Delta\nu$ results from the Heisenberg uncertainty principle, $\Delta\nu\Delta\tau \simeq 1$. In the oscillating maser, the narrow resonance line width is electrically equivalent to an oscillator having very high Q; this Q, designated Q_l in eq (1), is proportional to the storage time $\Delta\tau$. The net effective storage time, equal to the reciprocal

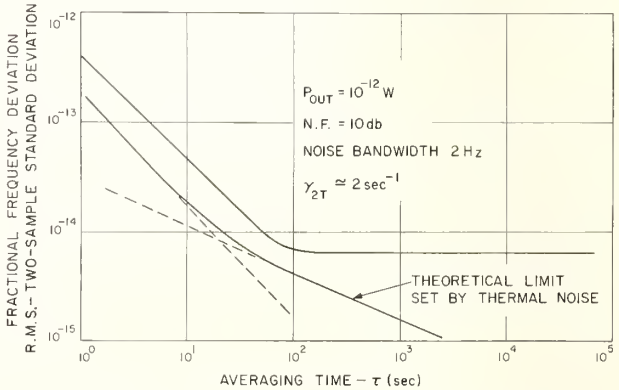


FIGURE 2. *Fractional frequency stability as a function of averaging time.*

of the net relaxation time γ_{2T}, results from a series of relaxation processes that go on in the storage bulb:

$$\gamma_{2T}=\gamma_b+\gamma_{2w}+\gamma_{2se}+\gamma_{2m}, \tag{2}$$

where $\gamma_b=\bar{v}A_aK/4V_b$ is the escape rate of the atoms from the bulb. Here $\bar{v}$ is the average velocity of the atoms, V_b is the bulb volume, A_a is the aperture area, and K is the collimator kappa factor for the entrance-exit hole [4]. Also, $\gamma_{2w}=(\bar{v}/\lambda)w_{2r}$ is the wall-collision relaxation rate, λ is the mean distance between collisions ($=4/3R$ in the case of a spherical bulb), and w_{2r} is the probability of a dephasing collision at the bulb wall. Chemical recombination of hydrogen or chemical interaction with the wall destroys the phase coherence of the atomic oscillating dipole moment with the rf magnetic field and is a dephasing collision.

The spin-exchange dephasing collision rate is $\gamma_{2se}=n\bar{v}_r\sigma$. It depends on the *relative* atomic velocity $\bar{v}_r$, the spin-exchange cross section σ, and the density n of the atoms in the bulb. The term γ_{2m} is the dephasing effect of the dc magnetic-field gradients in the bulb. The oscillation frequency of the coupled electron and proton magnetic-dipole moments has, in the "field independent" transition, a second-order magnetic-field dependence $\Delta\nu_m=2750\bar{H}^2$, where H is the magnetic field. The atom will gain or lose phase with the rf field as it moves in the bulb if the dc magnetic field varies from place to place within the bulb.

Magnetic gradients can also cause transitions among the $F=1$ magnetic sublevels at magnetic fields where the collision rate of the atoms is near the transition frequency of the sublevels.

The strongest limitation on the long-term stability encountered thus far is the effect of cavity mistuning. Since the maser is an active oscillator, the cavity resonance frequency affects the output frequency in proportion to the ratio of the cavity Q to the line Q. This "pulling" effect is described approximately by the expression

$$\Delta\nu_{\text{out}}\approx\Delta\nu_{\text{cavity}}(Q_c/Q_l). \tag{3}$$

In practice, there is a lower limit on the cavity Q set by the requirement that the maser oscillate. An oscillation parameter q, given by Klepper et el. [4], relates the cavity-bulb geometry, the cavity Q, and other constants. The expression for q is

$$q=\frac{\sigma\bar{v}_r\hbar\gamma_{2T}V_c}{8\pi\mu_0^2\gamma_b\eta V_b}\frac{1}{Q}\frac{I_{\text{tot}}}{I},$$

where

σ is the spin-exchange cross section for hydrogen,
$\bar{v}_r$ is the average relative velocity of the atoms,
$\hbar$ is Planck's constant divided by 2π,
μ_0 is the Bohr magneton,
V_c and V_b are the cavity volume and bulb volume, respectively,
Q is the loaded cavity quality factor,
η is $\langle H_z\rangle_{\text{bulb}}^2/\langle H^2\rangle_{\text{cavity}}$, and
I_{tot}/I is the ratio of the total atomic flux to the flux in the desired state.

Here, H_z is the Z component of the rf magnetic field.

For oscillation, the value of q must be less than 0.172.

The threshold flux for oscillation, neglecting spin-exchange processes, is

$$I_{\text{th}}=\hbar V_c\gamma_{2T}^2/4\pi\mu_0^2Q\eta.$$

The power delivered by the beam to the cavity[1] is

$$P=\tfrac{1}{2}(\hbar\omega)I_{\text{th}}[-2q^2(I/I_{\text{th}})^2+(1-3q)(I/I_{\text{th}})-1]. \tag{4a}$$

The maximum power is

$$P_{\max}=(\omega\hbar^2V_c\gamma_{2T}^2/64\pi\mu_0^2Q\eta)[1-(6/q)+(1/q^2)]. \tag{4b}$$

The flux at maximum power is

$$I_{\max}P=(\hbar V_c/4\pi\mu_0^2Q\eta)\gamma_{2T}^2[(1-3q)/4q^2]. \tag{4c}$$

Normally, the loaded cavity Q is about 35,000 and the line Q is about 1.5×10^9, making the ratio Q_c/Q_l about 2×10^{-5}. Thus, a shift of the cavity frequency of about 100 Hz will cause a change of 2×10^{-3} Hz in the output or a fractional frequency shift of about 1.5×10^{-12}.

There are several possible approaches for reducing the effect of cavity "pulling." One is to stabilize the cavity by improving the mechanical and thermal properties of the cavity. Another is to employ a cavity-tuning servosystem using a stable reference oscillator. This servo constantly maintains the cavity near its proper frequency by modulating the line Q and correlating the shifts in frequency with the modulation. The cavity is kept tuned by maintaining the condition that no correlation exist between the modulation and the frequency shifts.

A third approach, involving a fundamental feature of the hydrogen maser, is to extend further the storage time of the atoms in the bulb. At this point, it should be recognized that in the self-oscillating maser, the reduction of the atomic line width is accompanied by a reduction in output power. This results from the γ_{2T}^2 term in the quantity I_{th} in eq (4). The stability at short time intervals is controlled by the first term in eq (1), which gives the τ^{-1} behavior. The power depends on γ_{2T}^2, and the line width on γ_{2T}^{-1}. With $Q_l=\omega/2\gamma_{2T}$, the product of the first term in the brackets and $kT/2P$ is adversely affected and the second term is unaffected. The line width reduction helps directly in the cavity-pulling expression. The effect of varying γ_{2T} on the long- and short-term stability is shown in figure 3. For this family of curves, the maser receiver noise figure and bandwidth are kept constant, the cavity bulb geometry and Q are kept constant, and only γ_{2T} is varied. The contributions from a cavity-mistuning perturbation having a particular $1/f$ or "flicker" frequency spectral density are shown by the horizontal lines for $\sigma(\tau)$ [2], the contributions from additive white phase noise are shown by the τ^{-1} slope in the lines, and the constant limit set by thermal noise within the line is shown by a line with slope $\tau^{-1/2}$.

[1] If the wall-relaxation is due to recombination of atoms.

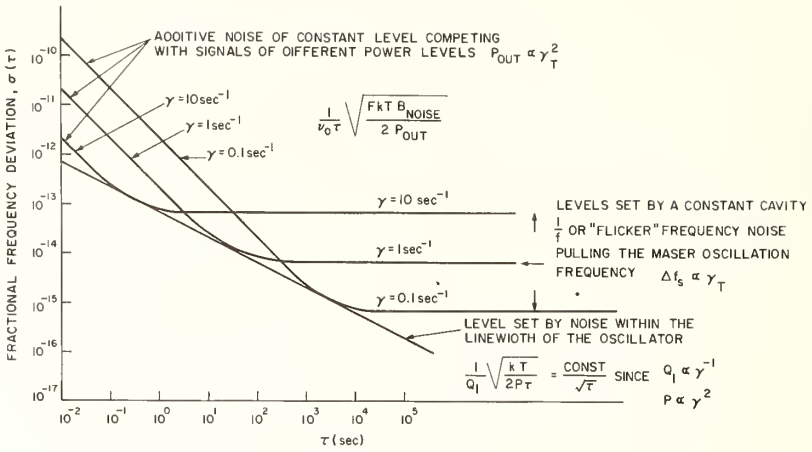


FIGURE 3. *Effect of varying the relaxation rate γ_{2T} in the maser:* $\gamma_1 > \gamma_2 > \gamma_3$.

1.3. Accuracy Limitations

The following four effects cause a bias in the frequency of the maser oscillator. They are systematic shifts, as distinguished from the random effects described above.

A. The magnetic-field behavior of the $F=1$, $m_F = 0 \rightarrow F=0$, $m_F=0$ transition shown in figure 1 is described by

$$\Delta\nu_m = 2750\langle H^2\rangle_b,$$

where $\Delta\nu_m$ is in hertz when H is in oersteds.

The squared magnetic field $\langle H^2\rangle_b$ is averaged over the volume of the bulb. To establish the value of $\Delta\nu_m$, the average value of H is determined by externally causing $\Delta m_F = \pm 1$ transitions among the $F=1$ state sublevels and observing the frequency ν_Z at which the maser output is quenched. The average field H in the bulb is obtained by the expression $\langle H\rangle_b = \nu_Z/1.4\times10^6$. Since ν_Z is easily measurable to better than 5 Hz, at the normal operating value of 10^{-3} Oe, the value of $\langle H\rangle_b$ is known to better than $\pm 3\times10^{-6}$ Oe and $\Delta\nu_m$ can be determined to 1 part in 10^{14}. If gradients are present in the bulb, it is not appropriate to use $\langle H\rangle_b^2$ in place of $\langle H^2\rangle_b$. However, the effect of magnetic gradients can be determined and minimized by observing the quenching of the maser through the γ_m term previously discussed. The error in $\Delta\nu_m$ resulting from taking $\langle H\rangle_b^2$ for $\langle H^2\rangle_b$ is about 1 part in 10^{14}, giving an overall error in $\Delta\nu_m$ of a few parts in 10^{14}.

B. The second-order Doppler frequency shift results from the thermal motion of the atoms within the bulb. Since the atoms on the average spend about 0.5 s in the bulb and make in the order of 10^4 collisions with the walls, the temperature of the bulb is taken as the temperature of the atoms. The second-order Doppler shift is given by

$$\Delta\nu_T = -\nu_D(3kT/2mc^2) = -1.9557\times10^{-4}T.$$

Here, k is Boltzmann's constant, T is in Kelvins, c is the velocity of light, and m is the mass of the hydrogen atom. Since we can measure the temperature to better than 0.5 K, the probable error in the second-order Doppler shift is in the order of 5 parts in 10^{14}.

C. The effect of cavity mistuning on the output frequency results from the pulling effect described earlier. There is also a systematic shift due to atom-atom collisions [5]. One can make use of the spin-exchange interatomic collisions to modulate the line Q of the maser for tuning purposes, as described earlier. When this is done, the output frequency is shifted by the pulling process and a systematic shift due to interatomic collisions. Both these processes are proportional to the line width and are described by the expression for the mistuning shift:

$$\Delta\nu_s = \left(\frac{\Delta\nu_{\text{cavity}}}{\nu_0} Q_c - \frac{13\sqrt{2}}{32}\frac{\bar{v}\hbar a_0^2 V_c}{Q_c\mu_0^2\eta V_b}\right)\frac{\gamma_{2T}}{\pi},$$

which is a more exact version of eq (3).

If we modulate γ_{2T} by varying the flux into the bulb, we can make the quantity in the brackets zero by adjusting $\Delta\nu_{\text{cavity}}$. The density-independent output frequency ($\Delta\nu_s = 0$) resulting from this tuning process in the maser effectively removes the collision shift [6].

The effect of the cavity shift on the accuracy of the maser frequency depends on the type of tuning system used. If the device is constantly being tuned by a servodevice, the value of $\Delta\nu_s$ will have a zero mean value and the value of the stability function $\sigma(\tau)$ will probably vary as $\tau^{-1/2}$ for large values of τ.

If the system is tuned and left alone, the accuracy of the output frequency will depend on the precision of the tuning and the stability of the cavity resonance frequency.

D. The strongest limitation on the accuracy of the maser is due to the fact that the atom, upon collision with the walls, suffers an average phase shift per collision. The collision process can be briefly described by a Van der Waal's potential function with the usual r^{-6} dipole-dipole attractive term and a r^{-12} term due to exchange forces. During its approach to the wall, the atom is polarized and experiences a r^{-6} type of potential, gaining energy until it en-

counters the r^{-12} repulsion due to close-range electron-exchange forces. It is therefore polarized during the time it is in the r^{-6} potential, resulting in a Stark-effect shift in the energy levels and a reduction in the ground-state hyperfine separation. The oscillating atomic dipole is thus retarded in phase with respect to the rf field in the cavity during the r^{-6} portion of its travel. During the repulsive portion of the collision, the electron cloud surrounding the proton can be visualized as being compressed, and during the time the r^{-12} forces are acting, the interaction energy between the proton and the electron is stronger, resulting in an increase in hyperfine separation. The oscillating atomic moment is thus advanced in phase during this part of the collision. It has been found [7] that the either positive or negative phase shift can predominate and that it is possible to establish conditions where an average zero phase shift per collision can be obtained.

The wall shift $\Delta\nu_w$ is usually determined by coating a number of bulbs of different sizes that provide a range of wall-collision rates that is as large as possible. The maser output frequency is then plotted as a function of collision rate. The value obtained on extrapolating to zero collision rate and correcting for magnetic and Doppler effects is taken as the unperturbed hydrogen frequency. The wall shift is related to the phase shift per collision as follows:

$$\Delta\nu_w = (\nu_c/2\pi)\bar{\phi}(T),$$

where ν_c, the average collision rate, equals $3\bar{v}/4R$ for a spherical bulb of radius R. The average phase shift per collision $\bar{\phi}(T)$ is a function of temperature, and $\bar{v} = (8kT/\pi m)^{1/2}$ is the average velocity.

Comparison of wall shifts is most often made by means of the quantity

$$K(T) = \Delta\nu_w D = 3\bar{v}\bar{\phi}(T)/4\pi,$$

where D is the bulb diameter.

2. Wall-Coating Materials and Techniques

The chief limitation in the accuracy of the maser as a frequency standard has been due to the wall shift. Accurate measurement of the wall shift is not easy and has depended on the ability of the experimenter to coat many bulbs in a reproducible manner. The value of the wall shift thus obtained can be applied with confidence only to the bulbs actually measured. Variations between laboratories in the measurement of this parameter have led to differences of some 5 parts in 10^{12} in the maser frequency.

In most cases, the material that has been used for coating the bulbs has been the FEP-120 Teflon dispersion made by duPont. This is a polytetrafluorethylene-hexafluoropropylene copolymer supplied in an aqueous dispersion; it is coated on the inside of the bulb, allowed to dry, and then sintered or fused to form a coating. Various coating processes have been used. However, all are essentially the same and involve burning off the wetting agent by heating in air and then heating to a temperature where the coating is fused to form a slick, clear coating. Adhesion to the inside of the bulb can be improved by roughing the fused silica surface with glass shot and abrasive grit until a ground-glass appearance is obtained.

Early measurements of the FEP material were made by Vanier and Vessot [8]. The wall shift that they gave for this material and that pertaining to the particular set of bulbs used in the measurement was until recently generally accepted as an intrinsic property of the material. This conclusion was wrong, since differences in the surface texture of the material can cause variations in the wall shift from one bulb to another.

Recently, there has been considerable activity in the study of the wall shift, using various types of Teflon bulb coatings. Zitzewitz [9] at Harvard found that Teflon showed a temperature-dependent shift that went from negative to positive, as shown in figure 4. Here, the data are given in terms of phase shift per collision versus temperature. The breaks in the curve, identified by α, β, and γ, are due to transitions in the internal short-range ordering of the molecules analogous to those in glass. Data are shown for two types of Teflon, the FEP copolymer and polytetrafluoroethylene homopolymer (PTFE duPont Teflon 42). The zero phase shift per collision temperature is about 107 °C.

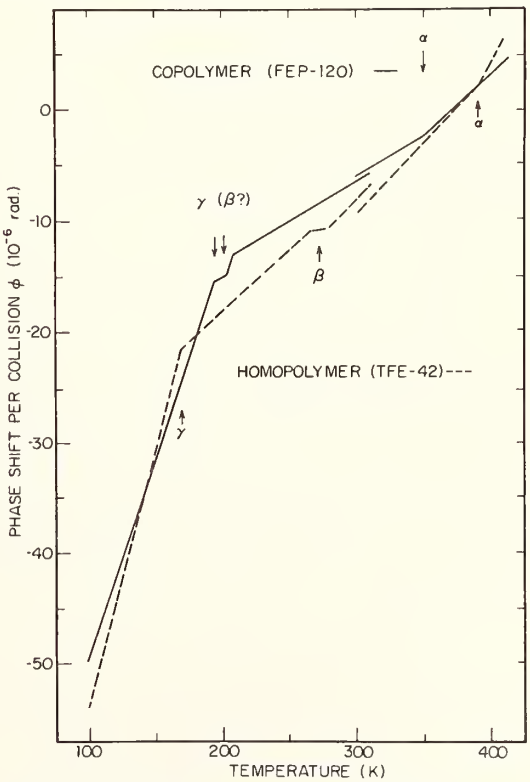


FIGURE 4. *Temperature dependence of phase shift per collision for PTFE and FEP Teflon.*

At the Smithsonian Astrophysical Observatory (SAO), studies on PTFE Teflon have been made by Vessot and Levine [10] using Teflon 42 under various conditions of curing. In these experiments, bulbs having integral collimators and smooth interior surfaces were coated with a single layer of Teflon dispersion. The dispersion consisted of 50 percent Teflon 42, 40 percent water, and 10 percent Triton x-100 (a wetting agent made by Rohm and Haas). The coatings were drained and dried for several hours by a slight flow of dry nitrogen introduced into the bulb through the collimator. The bulbs were sintered for 30 min at 360 °C in a horizontal air-flow oven with a constant flow of dry oxygen entering the bulb. The bulbs were then removed from the oven and quickly cooled in a stream of room-temperature air. It was found that the zero wall-shift temperature was reduced to 83 °C and that the wall-relaxation rate appeared to be very low. Very good reproducibility of phase shift per collision as a function of temperature was obtained for several bulbs.

The temperature slope of the phase shift per collision for the quickly cooled (quenched) PTFE was found to be 0.15 ± 0.004 μrad/°C. When the bulb was annealed in oxygen at a rate of 0.05 °C/min from 340 to 295 °C, the temperature slope decreased to 0.09 ± 0.01 μrad/°C and the intercept temperature went up to 142 °C. Zitzewitz gives values of 0.1294 and 0.1248 for the slope and a zero wall-shift temperature of 107 °C. For a recent experiment, he reports confirmation of the lower (83 °C) temperature at which the phase shift per collision is zero for a quenched PTFE-coated bulb.

With the same PTFE sample as that used by the SAO and Harvard groups, Hellwig at the National Bureau of Standards (NBS) made wall-shift measurements. His data taken at 24 °C gave a value of the wall-shift parameter K very close to the SAO value at the same temperature:

$$K_{\mathrm{NBS}}=0.528 \text{ Hz cm}$$
$$K_{\mathrm{SAO}}=0.522 \text{ Hz cm} \qquad \text{at } 24\ ^{\circ}\mathrm{C}.$$

To relate the values of the unperturbed hydrogen frequency among the laboratories at Harvard, SAO, and NBS and to give this frequency in terms of the ^{133}Cs hyperfine frequency that serves as the definition of the second, the SAO group undertook to make a traveling maser frequency-standard experiment. This work, reported by Hellwig et al. [11], gives the following new value for the unperturbed hydrogen hyperfine separation frequency in terms of the NBS cesium standard:

$$\nu_0=1\ 420\ 405\ 751.768\pm0.003 \text{ Hz}.$$

This value is in disagreement with that obtained by Vessot et al. [12] in 1965, who gave

$$\nu_0=1\ 420\ 405\ 751.7864\pm0.0017 \text{ Hz}.$$

The reasons for the difference are difficult to assess. Retrospective corrections for possible wall-shift errors in hydrogen and for possible errors in the frequency of the cesium devices have failed to close the gap between the two numbers and their error estimates.

In the present determination of the value of ν_0, the agreement in the unperturbed hydrogen frequency among all three laboratories is within ±0.002 Hz, or about ±1.4 parts in 10^{12}.

In summary, the PTFE wall-coating material appears to offer several advantages in performance over the FEP preparation:

A. The molecular weight of PTFE is larger by a factor of about 10 than that of FEP. The number of Teflon molecular-chain end groups for a given volume of material is thus reduced. The presence of atoms other than fluorine on the walls will cause different values of average phase shift per collision.

B. The relaxation rate associated with hydrogen collisions appears to be lower for PTFE than for FEP, especially between 50 and 100 °C. From preliminary measurements of relaxation rate, it appears possible to extend the storage time to about 5 s. This would substantially reduce the linewidth of the maser. The substance can be made highly crystalline, a property that reduces surface porosity and should further extend the storage time of the maser.

C. Structure does not change near the temperature at which the phase shift per collision is zero.

D. From recent data determined by different experimenters, wall-shift measurements appear more reproducible with PTFE than with FEP.

3. Elimination of the Wall Shift

The existence of a reasonably low operating temperature at which the wall shift goes to zero and at which the wall-relaxation rate is small makes possible a new method of measuring the wall shift.

Since the wall shift can be independent of the

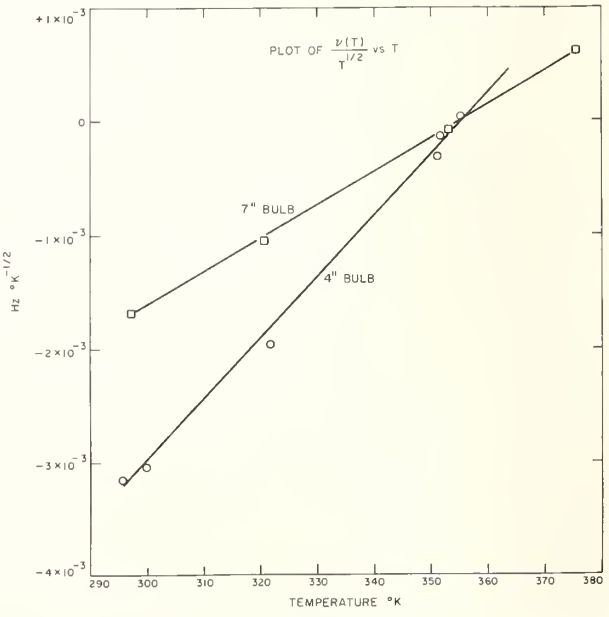


FIGURE 5. *Plot of* $\nu(T)/T^{1/2}$ vs T.

collision rate only when the average phase shift per collision is zero, it is not necessary to know the collision rate to obtain the zero shift temperature. One need only operate a maser with bulbs having different collision rates and determine the temperature at which the output frequency is the same. By obtaining the output frequency at this temperature the only further requirement for eliminating the wall shift is that the wall-coating material be the same for all bulbs.

Figure 5 shows wall-shift data taken at SAO [10] with a 4 in-diam bulb and a 7 in-diam bulb and plotted in terms of $\Delta\nu_w/T^{1/2}$. The division by $T^{1/2}$ removes the effect of the temperature dependence of the velocity of the atoms and makes the data linear and the intercept easier to compute. The intercept of the two resulting straight lines is well defined and gives a value of the unperturbed hydrogen frequency with an error $(0.2\pm1.1)\times10^{-3}$ Hz when related to the value obtained in the Harvard-NBS-SAO comparison reported earlier [11].

4. The Flexible Bulb

To overcome the uncertainties associated with the use of several separate bulbs to determine the wall shift, Brenner [13] suggested the use of a single flexible storage bulb that would allow the collision rate to be changed by a known amount without changing the nature of the surface. Recently, Debely [14] at Harvard conducted experiments on a maser using an open-end cylinder with a flexible cone at the other end that could be pushed into or pulled out of the cylinder to change the storage volume. This storage vessel is shown in figure 6. The movable part was made of a thin film of the FEP-120 copolymer that was also used to coat the bulb and the collimator. The ratio of maximum to minimum volume in this experiment was 1.299 ± 0.004. By tuning the maser in each of the positions, Debely obtained the wall shift of a reference maser, with an uncertainty of ±0.0023 Hz. In his design, he took pains to see that the annular joints of the Teflon coating of the bulb to the movable cone did not move or expose a different surface when the cone was moved.

The flexible-bulb technique can be used to establish the temperature at which the phase shift per collision is zero. This removes the uncertainty in the constancy of the wall shift with time and also overcomes the problem of using separate bulbs. By systematically changing the volume of the bulb (retuning the maser cavity at each position) and comparing the output frequency in each position against a stable frequency reference, one can adjust the temperature so that no change is observed in the output frequency. This procedure could be made automatic [15] by using servotechniques to maintain the proper temperature and promises to help greatly in ridding the maser of its most serious limitation on accuracy.

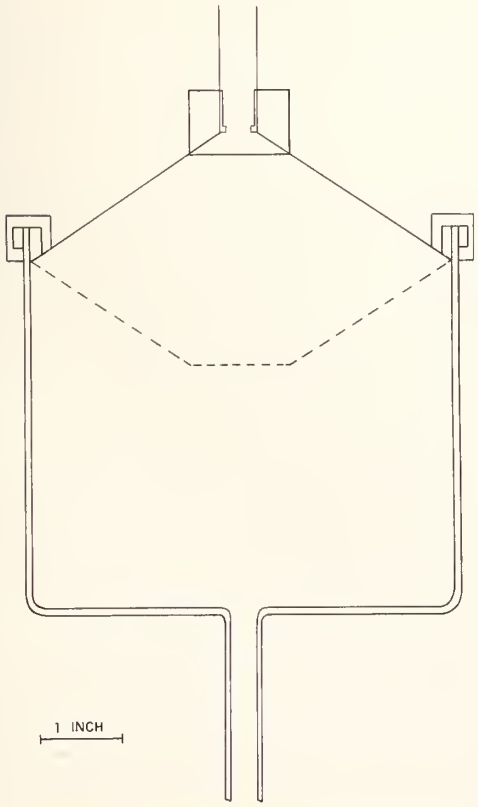


FIGURE 6. *Variable-collision-rate storage bulb.*

5. Multiple-Region Maser With Reduced Wall Shift

Since the wall-shift and relaxation rates in the maser depend on the collision rate of the atoms in the storage vessel, an important advantage can be gained by simply enlarging the storage bulb. However, limitations on bulb diameter of the self-oseillating maser are imposed by the cavity Q and the filling factor combination V_b/V_c that appears in the oscillating parameter q (see eq 4). A new concept of the maser was introduced in 1968 by Uzgiris and Ramsey [16] to overcome this limitation. A schematic of the new maser is shown in figure 7. The storage vessel consists of a 60 in Teflon-coated storage bulb connected to two rf interaction regions, each consisting of an rf cavity and a bulb, as in the usual maser. One cavity picks up the radiation from the atoms in its bulb, and this signal is amplified and sent to the other cavity to prestimulate the atoms in the other bulb. Since free molecule flow prevails in the three connected bulbs, the atoms can be put into a radiant state to emit at an enhanced rate in the low-level cavity. The loop is thus closed and the system will oscillate. The role of the amplifier is to make up for the lack of cavity Q and of the filling factor in the expression for q. The voltage gain R in the amplifier is given by the ratio of the rf magnetic field in the stimulating cavity to the rf field in the detecting cavity. The value of q for this maser is simply the value given earlier for the usual maser multiplied by $(1/R)(V_B/V_D)$, where V_B is the large

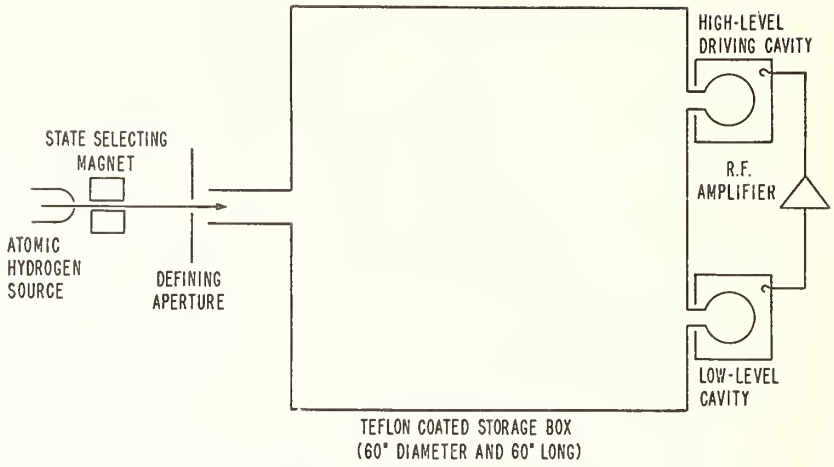


FIGURE 7. *Schematic of multiple-region hydrogen maser.*

bulb volume and V_D is the high-level cavity-bulb volume. The threshold flux I_{th} in this case is the usual value times a factor $1/R(V_B/V_D)^2$. Uzgiris calculates that the minimum gain required is about 70 dB and found that about 80 dB is required to make the maser oscillate.

The cavity pulling effect is complicated by the fact that both cavity resonant frequencies and the phase of the amplifier response must be included in the pulling expression along with a very small flow-dependent term that is not line width dependent. Except that the small effect of this term (in the order of 1 part in 10^{13}) must be accounted for, the tuning procedure is similar to that ordinarily used. The ratio of high-flux line-width to low-flux line width is about 1.7, as in conventional masers.

With the bulb geometry used in the experiment, the confinement time is calculated to be 10 s ($\gamma_b = 0.1$). At low-flux operation, a value of γ_{2T} of about 0.33 s^{-1} was observed, and the data indicate an excess relaxation rate of about 0.2 s^{-1} over the sum of the spin-exchange and bulb escape rates. This excess is most likely due to wall relaxation and is estimated to correspond to a probability of 10^{-4} per collision.

The results on the wall shift confirm the prediction that it would be 10 times lower than that encountered in the usual maser operating with the same wall materials and at the same temperature. A value of the shift of $(-2.3\pm0.8)\times10^{-3}$ Hz was obtained. At present, the multiple-region maser with a large storage vessel is being modified to incorporate the deformable bulb technique discussed in section 4 for measuring the wall shift of the maser coating that is actually in use. With this modification, the large-bulb maser should combine the advantage of a much smaller wall shift with the deformable technique for measuring that small shift.

6. Wall Relaxation

The storage-bulb technique depends on the ability of the bulb surface material to bounce the hydrogen atom without causing it to lose phase with the rf magnetic field in the cavity. Processes that lose the atom through chemical reactions or that change its atomic state and, thus, cause a loss of magnetization (T_1 processes), and collision processes that cause the oscillating atom to lose phase coherence with the rf field (T_2 processes) both contribute to the overall limitation on the storage time.

Early work by Berg [17] at Harvard gave values of the wall-relaxation probability w_r of about 1×10^{-5} per collision for FEP Teflon. Most of the masers constructed since then have used the FEP copolymer; and later measurements of the relaxation on this material [18] give values of w_r of about 1×10^{-4} per collision. The relaxation process has been found to be due mostly to recombination of the atomic hydrogen.

In recent work on PTFE, the value of relaxation probability per collision is approximately 1×10^{-5}, and the slow annealing process described earlier has yielded bulbs that give even more optimistic data. The difficulty in these recent results lies in the estimation of the bulb storage time, γ_b. Most of these bulbs have storage times $\sim$0.7 s ($\gamma_b\simeq1.4$ s^{-1}). The wall-relaxation rates have been estimated by measuring the exponential decay γ_{2T} of pulsed radiation after shutting off the atomic-hydrogen source and calculating $\gamma_w=\gamma_{2T}-\gamma_b$. For a 7 in bulb, values of $\gamma_w=0.1\pm0.1$ s^{-1} have been obtained. The sensitivity of the measurement can be improved by using bulbs with very long storage times, thus making wall relaxation the dominant effect. Work along these lines is in progress at SAO to determine if the annealed surfaces offer any improvement in the relaxation rate.

If the present data on PTFE are correct, it should be possible to extend the storage time by a factor of 3 to 5 beyond the limits obtained by using FEP Teflon. Values of $\gamma_{2T}=0.25$ s^{-1} or less should be possible in 7 in-diam bulbs. The benefits of this type of Teflon on the large bulb maser would also be very significant.

7. Electronic Systems

7.1. Cavity-Tuning Systems

a. Maser Cavity Resonator

The hydrogen masers constructed to date have all used a resonator operating in the cylindrical TE_{011} mode. This mode is convenient because, since only azimuthal wall currents flow, electrical contact between the cylinder and the end plates is not required.

Silver-coated CER-VIT[2] is used for cavity construction at SAO [19], while solid aluminum is preferred by the Goddard Space Flight Center [20]. The CER-VIT cavities are tuned by means of a voltage-variable capacitor mounted within the cavity and coupled to the rf magnetic field by a small loop. The aluminum cavities are thermally tuned; the relatively high temperature coefficient of expansion is used to advantage to permit tuning by small adjustments of cavity temperature.

b. Cavity Tuning

The oscillation frequency of the maser is pulled by tuning the resonator, as shown by eq (3).

There are essentially two problems associated with cavity pulling: (1) the cavity must be set to the correct frequency at the initiation of maser operation, and (2) the cavity must be held at this frequency indefinitely in the presence of environmental perturbations. The first problem is related to accuracy; the second, to stability. The setting of the cavity frequency can be accomplished by manual tuning techniques or by either of the two automatic tuners described in the next section. A series of careful experiments indicates that error in maser frequency attributable to cavity tuning is $\pm 7\times10^{-13}$ in a 90 percent confidence interval [11].

Once the maser is initially tuned, there remains the problem of maintaining this condition. One approach is to build the resonator out of materials of ultra-low thermal expansion, such as CER-VIT, which has a coefficient of expansion less than 5×10^{-8}/°C, and adequately control the temperature. Alternatively, the cavity can be servoed to resonance by continuous operation of an automatic tuning system.

c. Automatic Tuning Systems and Techniques

The hydrogen maser is tuned by modulating the line Q and adjusting the cavity resonator so that there are no correlated frequency shifts. Several automatic systems for timing masers have been devised; two of these are described below:

A. *Random-walk tuning* [20]. The masers used in this system are equipped with beam shutters so that the hydrogen flux and therefore the line Q can be modulated mechanically.

The maser frequency is compared to the reference oscillator, a stable crystal oscillator; a comparator circuit determines if the maser frequency was measured either high or low with respect to the reference. The comparator generates a "plus" pulse if the maser is higher in frequency than the reference and a "minus" pulse if it is lower. The plus and minus pulses are added algebraically in a reversible counter. When the comparator pulses accumulate to either a positive or a negative preset limit, a pulse of the appropriate sign is generated. These pulses are accumulated in a cavity-register circuit. A digital-to-analog converter translates the binary number stored in the cavity register to a cavity-control voltage that is used to servo the cavity to the correct frequency.

This random-walk automatic tuner, when used with an extremely low-noise crystal reference oscillator ($\sigma(\tau)=4\times10^{-13}$), will correct an initial maser-frequency error of 3×10^{-12} to approximately 2×10^{-13} in 20 h. The ultimate tuning error is claimed to be in the order of 5×10^{-14}.

B. *Fast automatic tuning* [19, 21]. The fast automatic tuner modulates the line Q by changing the rf power level into the hydrogen dissociator. The flux level must be modulated as rapidly as possible to minimize the phase-noise contribution from the reference oscillator.

The effect of linear frequency drift in the reference can be eliminated by using an algorithm devised by L. S. Cutler (private communication, 1968). The period of the beat between the maser and the reference is measured an odd number of times, alternately at low and at high flux. The measurements at low flux are assigned a negative sign; those at high flux, a positive sign. The first and the last measurements are divided by 2, and all the measurements are then summed algebraically. This residue,

$$N_R=-\tfrac{1}{2}(a_1+a_n)+\sum_{k=2}^{n-1}(-1)^k a_k,$$

is a measure of the cavity detuning and is independent of linear frequeney drifts in the reference oscillator.

In practice, each beat measurement is approximately 6.6 s, and a eomplete cycle consists of 21 measurements, 11 at low flux and 10 at high flux. The period measurements are taken on a reversible counter, counting down for the low-flux measurements and up for the high-flux. The addition of one-half the 1st and 21st measurements is accomplished automatically by dividing the clock rate by 2 for these two measurements. The binary number in the counter at the end of the cycle (the residue) is translated by a digital-to-analog converter into a voltage that is used to correct the cavity frequency.

The precision of this tuning procedure is approximately $\pm7\times10^{-13}$ after 2 to 3 h of operation. Normally, the maser is initially tuned and then left unperturbed except for periodic tuning checks. Continuous automatic tuning is usually not desired, because of the average reduction in line Q required by the modulation process. A further difficulty in continuous tuning can arise from the changes in maser output-signal level with flux modulation. Un-

[2] Owens-Illinois trade name for a partly vitrified glass with a low expansion coefficient.

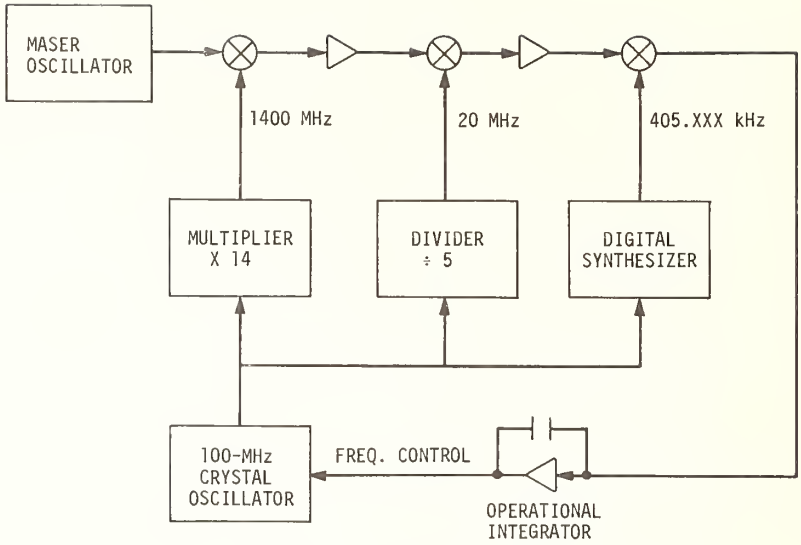


FIGURE 8. *Phase-lock system.*

less extraordinary precautions are taken in the design of the phase-lock circuitry, significant phase shifts will accompany amplitude changes. These phase shifts will not affect the automatic tuner, because they occur between operating cycles of the autotuner; but they will degrade the performance of external clocks.

Long-term frequency-stability data between two masers each operating with autotuning servos have been taken [22]. A value of $\sigma(\tau) = 2 \times 10^{-14}$ for $\tau =$ 12 h has been obtained.

7.2. Phase-Lock Systems

a. Description

The hydrogen maser is an active oscillator with an output of approximately −96 dBm at 1420.4 MHz. A crystal oscillator is phase-locked to the maser signal by a simple dual-conversion system, shown in figure 8. To provide standard frequencies at useful levels, the synthesizer is used to set the time scale.

The output of the maser at 1420.4 MHz is heterodyned in a low-noise balanced mixer with the output of the frequency multiplier at 1400 MHz. The 20.4-MHz output of the first mixer is amplified and heterodyned in the second mixer with a 20-MHz signal derived from the crystal oscillator. The resulting 405-kHz signal is amplified and used as one input to the phase detector. The reference signal is provided by the digital frequency synthesizer, which is referenced to the quartz oscillator. The dc output of the phase detector is used to phase-lock the quartz-crystal oscillator through a high-gain operational integrator. Double integration is incorporated into the phase-lock loop so that the phase-tracking error is very nearly zero. The time constant of the integrator is selected such that the closed-loop response has a bandwidth of about 20 Hz and is critically damped.

b. Recent Developments

Before 1968, the frequency of the crystal oscillator shown in figure 8 was usually 1 or 5 MHz. Recently, a high-quality 100-MHz quartz-crystal oscillator has been substituted for the lower frequency slave oscillator.

The use of the 100-MHz oscillator offers two important advantages. First, the additive phase noise at 1400 MHz is reduced by approximately 26 dB for the same oscillator-drive level. This permits the use of a much wider phase-lock-loop bandwidth and improves the overall signal-to-noise ratio. Typically, the 10- to 20-Hz bandwidth used with the 5-MHz slave oscillator can be increased to 100 to 150 Hz for the 100-MHz quartz oscillator. Second, the multiplication ratio to 1400 MHz is reduced from 280 to 14; this is of particular importance because the multiplier chain is a source of phase noise.

Usually, the phase noise is very small at 1400 MHz since at this frequency the local oscillator is most directly compared with the maser. The phase of the crystal oscillator, however, is steered to make up for phase changes in the multiplier, and if the output of the crystal oscillator is used, phase noise will be observed. Generally, the spectral density of the output-frequency fluctuations exhibits a flicker of frequency $(1/f)$ behavior. It has been observed that for a typical $\times 280$ multiplier driven by a 5-MHz oscillator, the $\sigma(\tau)$ function is unaffected at 1 s, but in the range from 1 to 100 s, there is degradation of stability. The high-frequency crystal oscillator substantially improves the performance in this range, and the value of σ (100 s) is limited by the cavity stability.

8. Conclusion

Present stability performance of the maser frequency-standard system exceeds that of all other

known frequency standards; however, the accuracy of the device has justifiably been open to question, mostly owing to problems involving the wall shift. The improvements that can be brought about as a result of recent developments when engineered into operational devices will make a great deal of difference in the acceptability of the hydrogen maser as both an accurate and a stable frequency standard.

9. References

[1] Goldenberg, H. M., Kleppner, D., and Ramsey, N. F., Phys. Rev. Letters **5**, 361 (1960); Kleppner, D., Goldenberg, H. M., and Ramsey, N. F., Phys. Rev. **126**, 603 (1962).

[2] Cutler, L. S., and Searle, C. L., Proc. IEEE **54**, 136 (1966).

[3] Allan, D. W., Proc. IEEE **54**, 221 (1966).

[4] Kleppner, D., Berg, H. C., Crampton, S. B., Ramsey, N. F., Vessot, R. F. C., Peters, H. E., and Vanier, J., Phys. Rev. **138**, A972 (1965).

[5] Bender, P. L., Phys. Rev. **132**, 2154 (1964).

[6] Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. Letters **11**, 338 (1963).

[7] Zitzewitz, P. W., Preliminary Report on the Temperature Dependence of the Wall Shift of the Hydrogen Maser for FEP and TFE Teflon, Harvard University, October 2 (1969).

[8] Vanier, J., and Vessot, R. F. C., Metrologia, **6**, 116 (1970).

[9] Zitzewitz, P. W., Ph. D. Thesis, Harvard University (1970).

[10] Vessot, R. F. C., and Levine, M. W., Metrologia, in press.

[11] Hellwig, H., Vessot, R. F. C., Levine, M. W., Zitzewitz, P. W., Allan, D. W., and Glaze, D. J., IEEE Trans. Instr. Meas., **IM-19** (November 1970).

[12] Vessot, R., Peters, H., Vanier, J., Beehler, R., Halford, D., Harrach, R., Allan, D., Glaze, D., Snyder, C., Barnes, J., Cutler, L., and Bodily, L., IEEE Trans. Instr. Meas. **IM-15**, 165 (1966).

[13] Brenner, D., Bull. Am. Phys. Soc. **14**, 943 (1969).

[14] Debely, P. E., Rev. Sci. Instr., in press.

[15] Smithsonian Astrophysical Observatory, New Technology Disclosure, NASA Contract NSR 09-015-098, June 1 (1970).

[16] Uzgiris, E. E., and Ramsey, N. F., Phys. Rev. **A1**, 429 (1970).

[17] Berg, H. C., Phys. Rev. **137**, 1621 (1965).

[18] Vanier, J., and Vessot, R. F. C., IEEE Trans. Instr. Meas. **IM-13**, 185 (1964).

[19] Levine, M. W., and Vessot, R. F. C., Radio Sci., **5**, 1287 (1970).

[20] Peters, H. E., McGunigal, T. E., and Johnson, E. H., Proceedings of the 22nd Annual Symposium on Frequency Control, U.S. Army Electronics Command, Ft. Monmouth, N.J., April (1968).

[21] Hellwig, H., and Pannaci, E., Proc. IEEE (Letters) **55**, 551 (1969).

[22] Risley, A. S., Allan, D. W., Peters, H. E., Johnson, E. H., Vessot, R. F. C., Levine, M. W., Gray, J. E., Shoaf, J. H., Machlan, H. E., and Glaze, D. J., Conference on Precision Electromagnetic Measurements, Boulder, Colo., June (1970).

LENGTH STANDARDS AND THE VELOCITY OF LIGHT

A stabilized laser
Competes with a maser,
Providing a standard of time.
Our unit of length
Does not have such strength
And isn't considered sublime.

ANON.

Length Standards

K. M. Baird

National Research Council Laboratories of Canada, Division of Physics, Montreal Road, Ottawa 7, Canada

This paper reviews the history of length standards from ancient times to the present from the point of view of putting into perspective new alternatives to the present international length standard, the 0.606 nm spectral line of ^{86}Kr. Possibilities of some of the new alternatives are discussed including the use of lasers stabilized by the technique of saturated absorption with reference to molecules such as I_2, CH_4, and CO_2 and the use of direct comparison of optical and microwave frequencies.

Key words: Distance measurement; length standards; stabilized lasers; standard wavelength.

1. Length Standards

Current developments in physics and technology are leading to the possibility of defining a new fundamental standard of length that will be orders of magnitude more accurate than the present meter standard. Furthermore it appears already that there may be a wide choice as to the best basis for the new standard. The adoption of a world standard is not an easy thing, involving as it does diplomatic and legal questions as well as sometimes difficult scientific ones; it is therefore important to approach with considerable care the crossroads we're coming to. It is the purpose of this talk to review the background of present day length measurement in order to present an overall perspective and to point out some important considerations that might get lost in the enthusiasm for the exciting new developments.

Let us start with a brief look at history, starting with very early examples of length standards, not just for entertainment but also to see what lessons there are for the advocates of various proposed new standards.

2. Some Ancient Standards

The cubit is the best known of early length units and was used in ancient Babylon, Egypt, Roman times and up into European medieval times. Cubit standards have been found on ancient structures and had lengths of about one-half a meter but varied by at least 10 percent from place to place and time to time. It was based originally on the length of the forearm.

FIGURE 1. *A sixteenth century length standard.*

About 5000 years ago the Chinese based a standard of length on the distance between knots of a piece of bamboo which, when used as a whistle, would emit a certain specified note, though how it was specified I don't know.

In Europe up to about 1800 A.D. most length standards were based on parts of the human body such as the foot. An old woodcut (fig. 1) illustrates the legal rod as being equal to the lengths of the left feet of 16 men as they left church on Sunday morning; an interesting example of statistical sampling and of specification of all the parameters that might be relevant.

The French Toise, equal to 6 "pieds du roi" was the principal unit of length in Paris up to the introduction of the metric system but, in spite of many royal attempts at unification, a variety of differing standards for this unit existed throughout France and Europe, adding to the difficulties of "honest commerce" and exact science on an international scale. A Toise standard was used in the measurements involved in setting up the metric system.

These examples illustrate a common attempt to have as a standard a "natural" or independently reproducible one, which however failed in the resultant use of practical measuring standards that were more precise in their application, but differed from one another by much more than their individual precisions.

3. The Metric System

The upheaval of the French Revolution provided in that country, the opportunity to introduce a new uniform system, having decimal subdivisions (corresponding to the number system) and having greatly increased accuracy as a result of the provision of a suitable standard. Again, following the urge for a natural standard, a choice was made between the length of a pendulum having a certain specified period, and one ten-millionth part of a quadrant of the earth's meridian. The former was rejected because it depended on "g" and therefore on geographical location and also because it was objectionable that the unit of time would enter into the definition of the standard of length. (This is interesting in view of present proposals we shall come to later.) So the 10^{-7} part of the earth's quadrant was chosen to be the standard and it was measured in terms of a Toise standard by a survey between Dunkirk and Barcelona, in spite of great political difficulties and personal danger due to the wars of the time.

The English decided not to cooperate, for their own reasons, and the Americans decided likewise, because the new unit was based on a measured meridian in France and Spain and was therefore not independently accessible. Thus the Anglophones were cursed with an antique system.

Table 1. *Measurement of the* Cd *red line with respect to the International Meter of 1889*
(in meters $\times 10^{-10}$)

Date	Observers	Corrected value in Normal air
1892	Michelson & Benoit	6438.4691
1905	Benoit, Fabry & Perot	6438.4703
1927	Watanabe & Imaizumi	6438.4682
1933	Sears & Barrell	6438.4713
1934	Sears & Barrell	6438.4709
1933	Kosters & Lampe	6438.4689
1934	Kosters & Lampe	6438.4690
1937	Kosters & Lampe	6438.4700
1942	Romanova, Varlich, Kartashev & Bartarchukova	6438.4677
1960	Baird & Hart	6438.4707

After the Meridian survey a practical standard was constructed of sintered platinum in the form of an end standard, i.e. one that defined a distance by the separation of accurately flat parallel ends. In 1799 this end standard, the Mètre des Archives, became in fact, though not by definition, the standard for the meter; it was actually 0.2 mm shorter than the Meridian definition would have required. It remained in force as the standard for nearly 100 years and was used as the basis for a new Meter which became a truly international one with the Convention du Mètre in 1875.

The new Meter standard defined a distance by the separation of two lines engraved on a bar of Pt-Ir (that is, it was a line standard); it was completed and compared to the Mètre des Archives in time to be adopted by the First General Conference of Weights and Measures in 1889 at which time equivalent copies were distributed to the participating countries. Careful measurements suggested that the new meter standard would be accurately reproducible to about 2 parts in 10^7 but it had sacrificed the characteristic of independent reproducibility in the interest of accuracy, demanding international cooperation in order to be effective.

During the nineteenth century Babinet (1829), Fizeau (1866), and Young had proposed that wavelengths of light be used as length standards. In retrospect, it is fortunate that the proposals were not taken up because the dubious advantage of having a natural standard would have been outweighed by the handicap of inaccuracy at a time when accuracy in length measurement was becoming very important to the progress of science and technology.

From 1889 to 1960 the Meter of 1889 served well as a world standard, providing the expected accuracy of $2:10^7$ which was good enough to satisfy any demands during that time. Evidence that its accuracy was maintained was provided by intercomparisons with other Meter bars, by comparison with end standards constructed from Brazilian crystal quartz and, most significantly, by intercomparisons with the wavelengths of light (the Cadmium red line at 644 nm). The first such comparison with optical wavelengths was made by Michelson and Benoit in 1892 (see table 1).

4. Wavelengths as Standards

The measurements of Michelson and Benoit, and succeeding comparisons of the meter with wavelengths, started a strong movement in favor of adopting an optical wavelength as the primary standard of length, and in fact the astronomers and spectroscopists adopted their own standard, the Angstrom, in 1922, defined in terms of the cadmium red line, in such a way that it was equal to 10^{-10} m to within experimental accuracy. During this time many thought that the wavelength standards were more accurately reproducible than the Meter itself but my impression of the evidence is that they were about equal, the major error in comparing the two arising from the process of comparison of two unlike standards. There were several unrecognized perturbations and unspecified relevant parameters which affected the Cd lamp emission. Also during this time, the technique of optical interferometry became highly developed and was used in the measurement of some geodetic bases and in the routine calibration of industrial gauges.

5. The Present Meter Standard

During the 1950's, following the development of single isotope spectral line sources, a number of lines were found and studied that were markedly superior as standards to the Cd red line. The result of this work and the development of interferometric techniques culiminated in the redefinition of the international Meter in 1960 in terms of the vacuum wavelength of the 606 nm line emitted by ^{86}Kr, which became the world primary standard of length (for the spectroscopists too, following the redefinition in 1961 of the Angstrom as exactly 10^{-10} m).

The new standard was specified in terms of the vacuum wavelength emitted by unperturbed ^{86}Kr atoms, but a source for practical metrology was recommended that would give the unperturbed wavelength to better than $1:10^8$, without regard to corrections. Extensive studies in a number of laboratories resulted in a quite precise knowledge of the perturbations due to ambient conditions such as discharge current, pressure, Doppler shifts and asymmetry in the line profile. When these are taken into account the accuracy of the standard can be improved by about an order of magnitude. (See for example fig. 2.) The width of the line at half intensity is about $8\times10^{-7}\lambda$. The practical form of the lamp is simple enough (fig. 3) so that the standard exhibits in an admirable way the desired attribute of making it possible to reproduce the international standard in a relatively modest laboratory, accurate enough for any requirement, up to the present.

A number of points connected with the work associated with the 1960 definition of the meter are worth

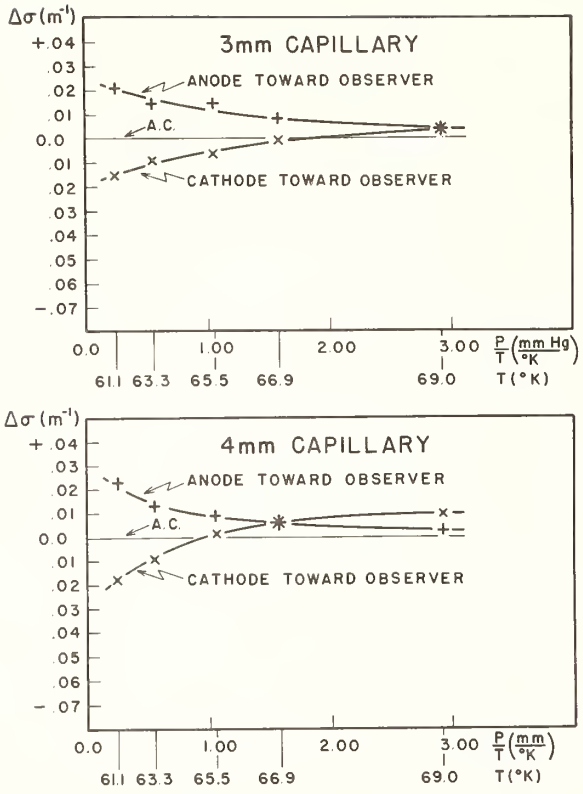


FIGURE 2. *Perturbations to* ^{86}Kr *standard caused by discharge current, direction of observation etc.*

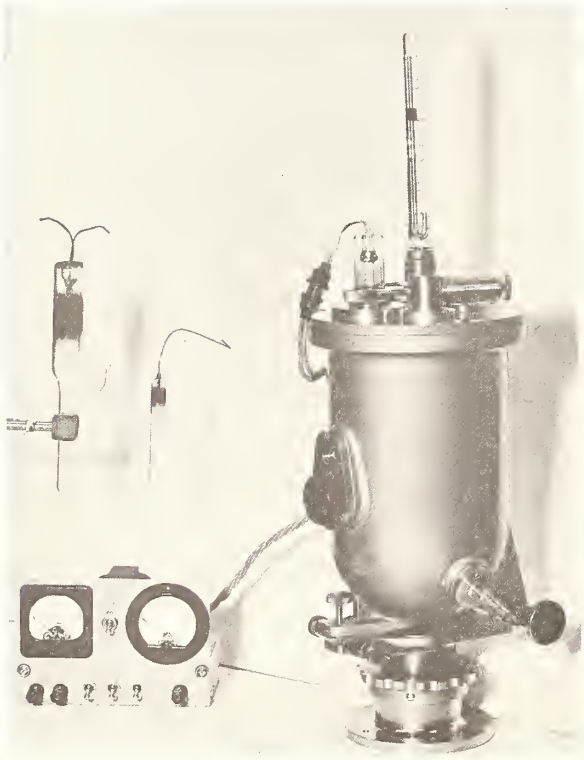


FIGURE 3. ^{86}Kr *lamp and cryostat housing.*

noting in the context of this paper: (1) Techniques were developed for setting on spectral lines and servo control of interferometers sensitive to about $10^{-10}\lambda$; (2) The accuracy figures of $10^{-8}\lambda$ and $10^{-9}\lambda$ are realistic and based on thorough and independently confirmed knowledge of perturbations; if two lamps constructed in the same laboratory were run side by side, as tends to be the criterion often used in early examination of potential new standards, it is likely that the differences between them would be found to be not much greater than $10^{-10}\lambda$ and in earlier investigations it was suggested that such accuracy would result with the ^{86}Kr standard; (3) There was considerable pressure at that time to adopt a wavelength standard for the Meter before extensive studies were completed but events have, without exception, proven that the conservative approach was well taken: one illustration is the hasty adoption of a wavelength for ^{198}Hg that was subsequently shown to be in error by $4:10^8$; another is the proposal to adopt an I_2 absorption line as a standard. The arguments for the latter were very good and in a sense anticipated action which will likely be taken in the future. However the I_2 "line" proposed then is really a large number of components spread in an unsymmetrical fashion over a range of about $10^{-6}\lambda$. These illustrate the foolhardiness of sanctioning a natural constant as a standard before it is thoroughly known.

6. The Laser

At the time of the adoption of the present meter standard, the laser had already been born and was soon being advocated as a superior replacement for the ^{86}Kr standard because its coherence and high power made very sensitive interferometry possible. However the laser itself does not emit a precisely fixed wavelength but one that can vary over a large part of the Doppler width of the line involved and it provides a standard only if stabilized by some reference such as a feature of the spectral line profile.

TABLE 2. *Early Measurements of Stabilized* He–Ne *Lasers*

Date	Laboratory	Laser	Vacuum wavelength (Meters$\times 10^{-10}$)
1963	NPL	----------	6329.9141
1964	NBS	----------	.9146
1965	NBS	----------	.9145
1965	NBS	Spect.-Phys.	.9147
1966	NPL	----do----------	.9138
1966	NPL	Elliot	.9145
1966	PTB	Spect.-Phys.	.9138
		after 1500 hours	.9134
1967	U.S.S.R	----------	.9145
1967	NRC	----------	.9148
1967	NBS	Same Spect.-Phys.	.9142
1967	NPL		$.9141_5$
1967	PTB		$.9141_6$

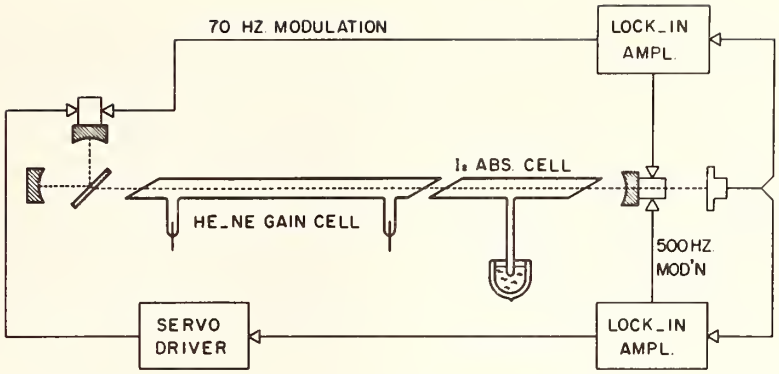


FIGURE 4. *Diagrammatic view of* I_2 *stabilized laser system.*

The lines used in early stabilized lasers were associated with atoms in an electrical discharge and were therefore not as good as the ^{86}Kr standard, which had been selected after extensive searching and is excited under conditions particularly well controlled to avoid perturbations and not having to be compatible with laser action. Such lasers have been much improved but in early stages were about as accurate as the Meter of 1889 (table 2).

The laser does, nevertheless, provide the tool to make possible the use of natural reference constants which would not otherwise be accessible, viz., spectral absorption lines uncluttered by Doppler broadening.

The Doppler broadening is eliminated by a velocity filtering effect: when an absorbing material is placed in an optical cavity, so that it is exposed to running waves propagated in both directions, some of the atoms or molecules will have a velocity component in the direction of propagation close enough to zero so that the Doppler shift is within the natural width of the line, and these atoms or molecules will interact with twice the intensity of radiation as others. Thus under conditions favoring saturation, there will be a reduction in absorption in this region centered about the absorption lines, and a laser line scanning over the power curve will have a net increased output corresponding to the natural line profile.

The arrangement of an I_2 stabilized laser at NRC is shown in figures 4 and 5 in which are shown the mode selector, the laser gain tube and the iodine absorption tube with its vapor pressure controller. Figure 6 shows the appearance of the derivatives of the peaks on the laser system output curve, caused

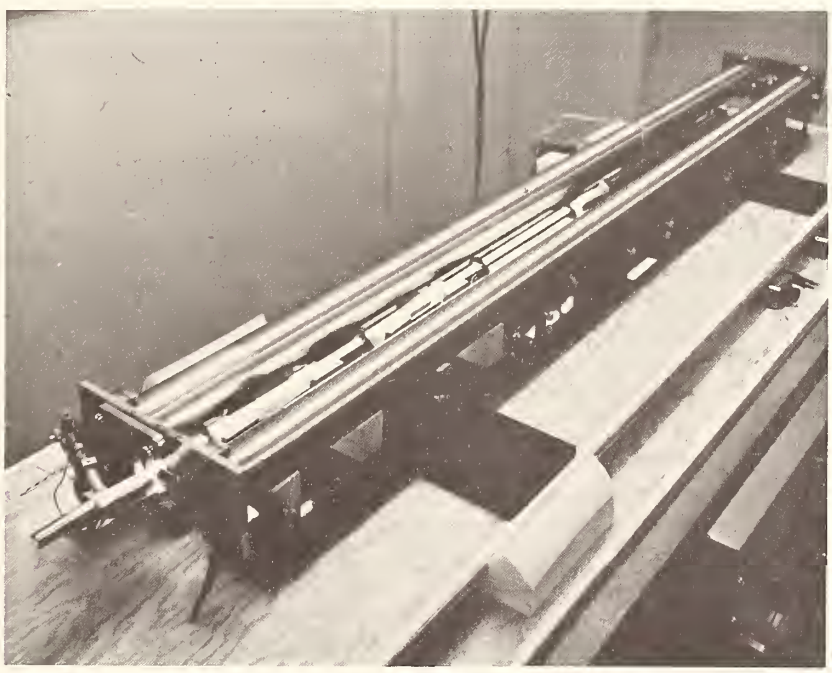


FIGURE 5. I_2 *stabilized laser system.*

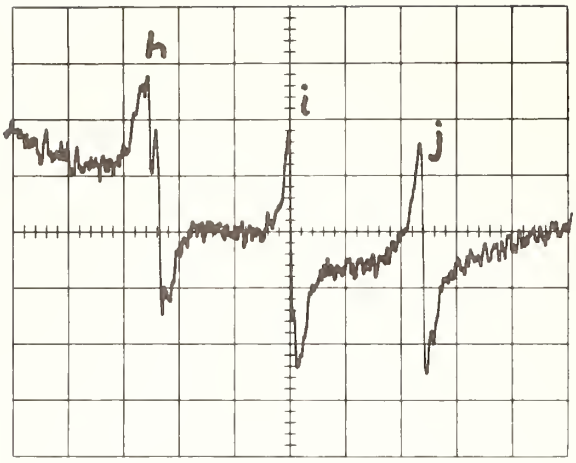


FIGURE 6. *Components of I_2 absorption line at 633 nm. (derivative).*

by saturated absorption of the I_2 hyperfine components and figure 7 indicates the location of the components with respect to the Ne gain curve and "Lamb dip". These lines at 633 nm have a width of about $9\times10^{-9}\lambda$ (4 MHz), perturbations due to controlled pressure and magnetic fields less than $10^{-12}\lambda$ (extrapolated from measured effects at high values of these parameters) and they can be pointed on to about the same value.

The first saturated absorption stabilized laser was the He-Ne-CH_4 system emitting at 3.39 μm developed at J.I.L.A. in Boulder, Colorado. The line width is about $3\times10^{-9}\lambda$ and perturbations appear so far to be controllable to at least $10^{-12}\lambda$. Independently run lasers of this type show relative deviation of less than $1:10^{11}$.

The CH_4 system has certain advantages over the I_2 system in that it is inherently quieter, has better signal to noise by virtue of a higher absorption peak and the absorption line is narrower. On the other hand the fact that the I_2 system is in the visible is a great convenience from the point of view of alignment, optics and detection.

There are a large number of other systems possible, among them the CO_2 laser in association with either CO_2 or some other molecule such as SF_6 as an absorber. I have not had confirmation but have heard by word of mouth that French workers have observed widths as fine as $10^{-10}\lambda$. *Note added in proof:* The author has since discovered that this was a false rumor arising from a confusion between line width and pointing accuracy. The line widths observed were about $10^{-8}\lambda$.

These stabilized laser systems require the finding of a coincidence between a suitable laser line and an absorption line, but the development of parametric emitters, tuneable over very great ranges, will open up an enormous number of possibilities as to the reference line, including atomic absorption lines. However the very fact that the emitter is tuneable over a wide range may militate against the desirable characteristic of a precisely controllable output frequency.

In summing up this development we note first that the much increased sharpness of the reference line, two or three orders of magnitude better than the present primary standard, plus the advantages of laser power and coherence, make possible a corresponding increase in pointing on the standard. This, in addition to the per se improvement in potential accuracy, makes possible a corresponding improvement in the study of perturbations. Secondly, because the reference line is an absorption line without a perturbing electrical discharge, and because it can be usefully employed at a lower pressure, the perturbations will be much less, as we have already found. Consequently, it is almost certain that a new

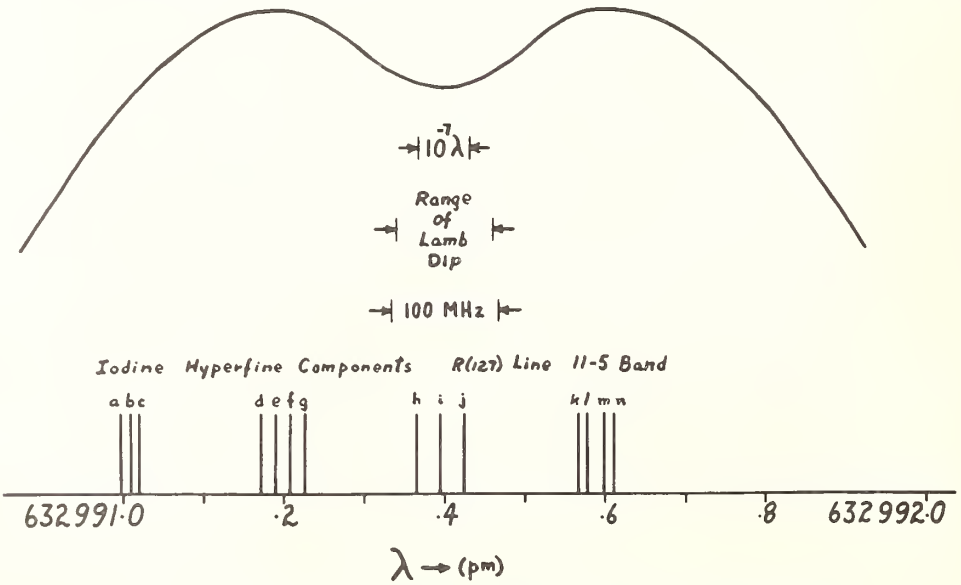


FIGURE 7. *Location of I_2 hyperfine components relative to Ne line at 633 nm.*

standard will prove after investigation to be absolutely reproducible to at least $10^{-11}\lambda$, as suggested by evidence so far.

It is interesting to compare these characteristics with those of the present time standards such as the Cs beam (width $= 10^{-8}\lambda$) and H maser (width $= 10^{-9}\lambda$). There seems to be some likelihood that standards in the visible or near I.R. will soon prove at least as precise as the present time standard.

A second way by which lasers have opened up the way to a redefinition of the fundamental length standard is that by virtue of their properties of coherence and power, they have made possible the direct comparison of frequency in the optical region with frequency standards in the microwave region. Since the velocity of light in vacuum is presumed to be an invariable natural constant, this means that vacuum wavelength in the optical region could be deduced from frequency standards if a conventional value for "c" were adopted. In such a case lengths would be measured in terms of a time standard, either through vacuum wavelengths or time of flight of pulses or modulation waves. The originators of the metric system would be appalled!

Alternatives to the direct comparison of frequency over this long range (a factor of 10^4) are provided by various velocity of light measurements now in the offing, some of which you will hear discussed later in this conference.

In order that such a course make any sense, the relating of frequencies in the optical range (whose wavelengths are required for precise short distance measurement) must be at least as accurate as the frequencies at each end of the range. (Even then the value of such a course of action could be questioned.) It should also be pointed out that from the viewpoint of the present it is not clear which frequency, the optical or the microwave, would be eliminated as redundant and insufficiently precise.

7. Present Accuracy of Length Standards and Measurements

Before going into what conclusions or lessons ought to be drawn from the evidence of the past, let us summarize the situation at present regarding length measurement requirements, techniques and standards. This is done conveniently by reference to figure 8.

At the right side are indicated the accuracies of various standards: The Meter of 1889 M(1889); the Engelhard lamp used in practical metrology, M(1960 prac); the ^{86}Kr standard M(1960) and the time standard (t). In addition are marked a lower limit of the (unconfirmed) reproducibility of saturated absorption lasers, L, the present certainty of the velocity of light, c(1970), and various other limits inherent in some distance measuring techniques, such as the refractive index of air corrected at the end only of a long path, n, and corrected for the total length of path by a dispersion method, n(corr). The letters at the left side of the diagram indicate the

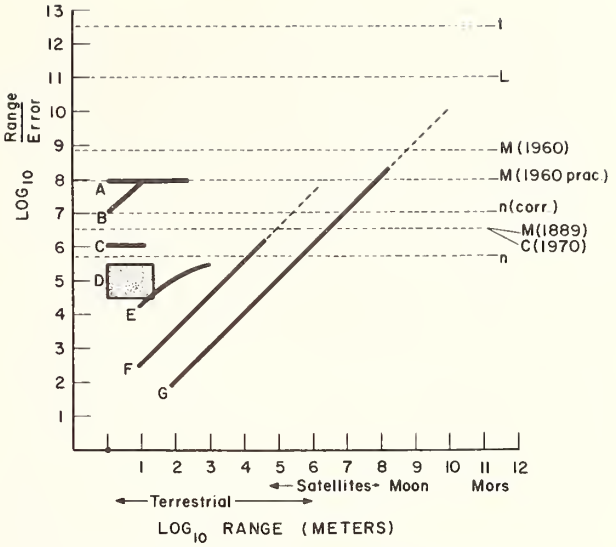


FIGURE 8. *Accuracy of various length standards and measuring techniques.*

accuracy of various measuring techniques, usually limited by some uncertainty inherent in the distance measured, such as that caused by irregularities in the engraved lines of a scale, for instance, or secular instability in the physical length of a material secondary standard. The letters designate the following techniques: A, optical interferometry; B, the best calibration of precise line standards (meter bars) and end standards; C, routine calibration of gauge blocks, geodetic tapes, etc.; D, engineering and manufacturing practice; E, ranging up to 1 km or so by Pockel's cell devices (such as the NPL mekometer) and modulated semiconductor emitters; E, long distance surveying by means of modulated microwave and optical wave instruments; F, radar ranging and space ranging.

The accuracies shown are of course only general indications which require qualification in most cases. For example the value of 10^8 for optical interferometry applies to careful wavelength comparisons and to the careful measurement of the displacement of optical surfaces. The displacement is usually very difficult to relate with this accuracy to some other type of physical length such as the length of a bar or base line, although it is done in some long and careful experiments. Usually a much lower limit is inherent in the nature of the distance being measured. On the other hand, where the distance is very precisely fixed such as in the current experiments on the Earth's crustal movements, interferometry would make possible a very much higher accuracy except for the limitation of the source and it is in this work that the new absorption line stabilized lasers find early application.

The relatively long range measuring techniques, D, E, F, depend on the velocity of light and in addition to fixed errors, due to resolution, are limited by uncertainty in the refractive index of air as well as,

for absolute measurement in terms of the international meter, by uncertainty in the knowledge of c, the velocity of light. If correction for the refraction of air is made by measurement only at the end point of a path of more than several hundred meters, the error is likely to be greater than $1:10^6$. By the use of a technique developed in the western U.S. whereby the refractive index along the whole path is inferred by a dispersion measurement, the errors can be reduced to the order of $1:10^7$.

It is difficult to imagine much improvement regarding uncertainty due to refractive index in measurement of long distances through the atmosphere. Therefore the knowledge of the value of "c" in terms of the meter need be improved very little to satisfy the requirements of terrestrial distance measurements of more than a few hundred meters. Shorter distances can be done by interferometry in terms of vacuum wavelengths.

On the other hand, in the case of extraterrestrial distances such as satellite or lunar ranging, the refractive index correction and the resolution are small compared with the range so that the distance can be measured in terms of time of flight of light to a much greater accuracy than "c" is known in terms of the meter. The obvious conclusion is that extraterrestrial distances should be given in such terms. Interstellar distances already are. One cannot help recalling the satisfactory duality of length standards, the meter and the angstrom, whose relation was known (10^{10} Å = 1 m) very accurately but not related by definition. There was no conflict as the Å was used for a special set of precise measurements involving optical and very short wavelengths.

As shown in the diagram practically all terrestrial length measurements are limited in accuracy to much less than the present ^{86}Kr wavelength standard. An important set of measurements are at present contemplated or already under way for the measurement of the velocity of light to a very much improved accuracy. An improvement in accuracy over the present value of two orders (not an easy thing to do) would still find the present standard adequate but there are high hopes of a much greater increase in accuracy than two orders and the success of the stabilized lasers as improved standards will be most welcome.

8. Conclusions

Let me conclude by drawing attention to two important points which strike me on considering the important task of changing the primary standard of length.

The first point is that, because a length standard is for the purpose of making possible uniform, accurate length measurement in a large variety of ways, the standard must be one that in practice is accurate, independently available in relatively modest laboratories and preferably reproducible in terms of a natural physical constant which satisfies these requirements. Philosophical attractiveness is of secondary importance and the exaggeration of its importance is one of the pitfalls that has been frequently fallen into. Three notable examples of the latter are the definition of the meter in terms of 10^{-7} of the Earth's quadrant, the use of "absolute" electrical standards which were a tenth as precisely reproducible as the working standards and the standard of light, the candela, which suffered the same drawback.

Secondly, let me plead the case that redefining a standard or international sanctioning of a value for a constant is not to be lightly undertaken; it ought to follow thorough investigation and confirmation in a number of laboratories and not be unduly pressed simply because an enthusiastic advocate has found that two similar experiments in his laboratory have produced very satisfying agreement. Published results of an experiment can always be used on their merits but official sanction of inadequately tested results can be a great nuisance. One would not have thought that this point would have to be made to scientists but I could give many examples where overenthusiastic pressure for following such a course has been applied, and even cases of the sanctioning of wrong values. The conservative approach has, in the past, been justified and the cases where there was scientific loss because of slowness in official adoption of new standards are rare indeed.

I ought to add that in spite of a conservative attitude, I am extremely optimistic about the possibility of the absorption line stabilized lasers and I think it quite likely that they will be proven at least equal to time standards in accuracy in the not too distant future.

I should like to close by giving my choice and that of my colleagues in Ottawa, as to a good bet for the measurement system for the future.

The CO_2 laser system with emission at 9 μm and 10 μm appears to us an unusual combination of good features:

(1) The laser is simple, of large cross section, high gain and single mode power.

(2) It appears so far to be very favorable from the point of view of stabilization by the saturated absorption technique either with respect to CO_2 lines or those of other molecules such as SF_6.

(3) The wavelengths at 9 μm and 10 μm are long enough so that short term fluctuations in interferometrically measured distance, due to vibrations, etc., are controllable to reasonable amounts relative to the fringe length and fringe counting rates will be low for a given rate covering distance; on the other hand the wavelength is short enough that diffraction corrections are not too great and that resolution of 10^{-3} of a fringe (0.005 μm) is sufficient for most purposes.

(4) The available choice of several transitions differing by 1 percent to 10 percent makes possible the use of the method of exact fractions, (such as is used in surveying distance measuring devices) to infer the whole number of fringes in the interference path, thereby avoiding counting over great distances.

(5) The spectral region appears favorable from the point of view of direct frequency comparison with microwave frequency standards.

The present disadvantage of difficulty in detection appears on the verge of being overcome (for example, with pyroelectric detectors or up-conversion to the visible).

Thus this system appears very attractive to us in having the possibility of providing at once a very accurate standard and one which will be very useful directly in practical distance measurement.

It is a very exciting and agreeable time to be associated with length metrology!

9. Bibliography

The historical references are taken from the Encyclopedia Britannica; Candler, C., Modern Interferometers (Hilger & Watts, London, 1951); and Moreau, H., J. Chem. Ed., pp 3–20, (Jan. 1953).

Work connected with the redefinition of the meter in 1960 in terms of optical wavelengths is covered in the proceedings of the Comité Int. Poids et Mesures for the period 1950–60.

Saturated absorption stabilized lasers are described by Hall and Barger, Phys. Rev. Lett. **22**, 4, 1969 and Hanes and Dahlstrom, App. Phys. Lett. **14**, 362, 1969.

Distance measurement techniques and accuracies are reviewed by Baird, Metrologia, **4**, 135, 1968.

DISCUSSION

D. Halford: I applaud the direction of considering the use of the iodine-saturated absorption to control a laser, and instead of being mortally afraid of it, actually we welcome this because there is a chance that that will be our next frequency standard, although there is an even better chance that some other saturated absorption device in the infrared or visible could be the next frequency standard. And, of course, there's still another chance that the best frequency standard would remain in the microwave or in the millimeter. But the primary thing is that the iodine saturated absorption device or the methane saturated absorption devices are indeed frequency standards first and foremost. That is how they are designed. That is how they work internally. And I recommend them to the length-measuring fraternity. (*Laughter.*) They are indeed recognizing that the frequency approach is a very useful way to do metrology.

K. M. Baird: Could I give a little answer to that? First of all, in measuring lengths we like to have a standard of length. Secondly, since I think you as much or more than anyone would admit that since the velocity of light can be considered invariant and therefore the relation between frequency and vacuum wavelength is fixed, you can say that either one of them is the standard. You don't have to say that it's a frequency standard any more than it's a length standard in vacuum if you accept the velocity of light as invariant.

D. Halford: I would like to agree with that, that indeed it is both. However, just to get to the point, when you build a servomechanism which controls the laser to the iodine absorption, it is not the wavelength of the laser which is stretched out and servoed to be equal to the wavelength of the iodine. Instead, it is the frequency which is servoed to be equal to the frequency of the absorption. If one were to build this inside by a length design, it would not be so good. It would be rather good but not so good.

A. Javan: May I take the privilege of being the chairman and make a very quick comment here in this connection? You see, after all, I think what we are talking about is whether two different frequencies can be compared with better accuracy, or we can compare two wavelengths using fringe-counting techniques. Which one can one do with better accuracy? I think we all agree that if you have two wavelengths that happen to be close, you can compare them with a high degree of accuracy if they are stabilized on the peak of the Lamb dip or whatever, maybe to a part in 10^{10}, possibly to a part in 10^{11}. But again, suppose you wish to compare two wavelengths, one let us say in the far infrared or in the middle infrared with another one in the infrared, using an interferometric technique. In my own experience at M.I.T., there is an experiment where we are attempting to compare a 10-micron laser in a wavelength measurement, an interferometric measurement, to another one at 6328 Å and a krypton standard. We are just about barely hoping to do it to a part in 10^8, maybe 10^9. How much better you can do it is hard to tell. But through this harmonic frequency mixing technique that we have I think two widely different frequencies can be compared. Do you agree on this point?

K. M. Baird: I think probably you're right. You see, the thing is though, to make this war fair, we should agree where the battle lines are drawn. And when you come up to the near infrared and the visible, that's our side. (*Laughter and applause.*)

Z. Bay: I would like to add one more remark to that question of wavelength and frequencies. They are even in vacuum not strictly connected by the velocity of light, because the phase velocity of light depends on diffraction.

A. Javan: You're quite right, yes.

Z. Bay: So you don't have simply the periodicity in space equivalent to periodicity in time.

A. Javan: I think the important consideration I mentioned on the limitation has to do with diffraction. I think that is true.

A. H. Cook: May I add another point to that? What is fundamental in the definition of the standard is the energy difference between atomic levels, and that is connected in a rather fundamental way to frequencies.

A. Javan: That's right.

A. H. Cook: And another comment to reinforce Dr. Baird's remark about being cautious: It was on

his slide but he didn't draw it out. Almost the very year in which the new mechanical definition of the meter was drawn up, Michelson showed how it was going to be superseded. And maybe someone is doing something in his laboratory now that will supersede something here.

A. JAVAN: Frankly, we are going back to this technique that we said initially was proved out in measurement of the lamp by measuring the period of the pendulum, so that is a perfectly good way of doing it. (*Laughter.*)

K. M. BAIRD: It wasn't then.

B. JAVAN: I have the prerogative of the chairman. (*Laughter.*)

K. M. BAIRD: It wasn't then. Yes, they should have had someone for our side. (*Laughter.*)

A. JAVAN: I really don't think there is any major argument on the subject.

The Implication of Saturated Molecular Absorption for the Laser Wavelength Standard Problem

J. L. Hall* and R. L. Barger**

Joint Institute for Laboratory Astrophysics, Boulder, Colo. 80302

The existence of very important pressure-dependent wavelength shifts of self-stabilized lasers has prevented the wavelength reproducibility of these lasers from substantially exceeding that of the krypton primary standard. The use of laser lines stabilized to accidentally-coincident molecular transitions was suggested as a conceptual means for securing laser optical advantages along with high wavelength reproducibility [1]. Almost immediately workers in several laboratories exploited this molecular reference concept by use of laser saturation of the absorption (the so-called "inverse Lamb-dip" effect), thus obtaining *both* extremely *sharp* and very *stable* reference transitions. The resulting sharp saturated molecular absorption resonances have enabled many groups to obtain unprecedented laser stabilities at 10.6μm, 3.39μm, and 0.633μm.

For example, with suitable He–Ne lasers at 3.39μ, using intracavity absorption cells containing methane, we have obtained optical resonance linewidths of less than 30 kHz. Servo electronics have been constructed which can pick the line center with satisfying precision. The resulting frequency stability improves approximately as $\tau^{1/2}$, where τ is the averaging time. Our present results show a probable fractional frequency variance of 3×10^{-4} at 1 ms, 1×10^{-12} at 1 s, and 1×10^{-13} for times of a few hundred to a few thousand seconds. The frequency stability does not continue to improve for longer averaging times due to *changes* in the same systematic effects which limit the present demonstrated *reproducibility* to $\leq$1 part in 10^{11}. Present efforts are directed toward engineering the major quantum jump in *accuracy* that systems analysis suggests is available.

It may be expected that several such laser-molecule systems will be found of roughly comparable quality. We are probably already past the level where improvements can be identified on the basis of *wavelength measurements*—even with our evacuated 30m high finesse Fabry-Perot interferometer. Nontechnical issues such as convenience and/or simplicity may well determine which particular laser system is best for a given project.

Based on the recent dramatic progress in IR laser frequency measurements [most recently up to 10.6μm (28 THz)] [2], it may be expected that all of these secondary wavelength standards will be compared directly to the primary cesium *frequency*. Thus a defined value for the velocity of light and, potentially, a single standard of both wavelength and frequency are concepts which underlie much of the present research effort.

Key words: Lasers; saturated, molecular absorption; wavelength standard.

---

* Staff member, Laboratory Astrophysics Division, National Bureau of Standards, Boulder Colo. 80302,
** Staff member, Radio Physics Division, National Bureau of Standards, Boulder, Colo. 80302.

References

[1] Baird, K. M., NRC, private communication; Zare, R. N., Columbia University, private communication; Gould, G., Polytechnic Institute of Brooklyn, private communication; Chebotayev V. P., et al., IEEE J. Quantum Electron **4**, 868 (1968); Shimoda, K., Proc. of Conference on Precision Electromagnetic Measurements, Boulder (1968).
[2] Evenson, K. M., Wells, J. S., and Matarrese, L. M., Appl. Phys. Letters **16**, 251 (1970).

DISCUSSION

A. Javan: With the saturated resonance I would like to point out that in CO_2 one has been able to observe this.

J. L. Hall: I have to comment on this.

A. Javan: Yes. The previous speaker mentioned something about the French having—there are rumors. What are the rumors? Maybe Dr. Baird could make some comment.

K. M. Baird: We had a visit a couple of weeks ago from someone from—I don't remember the institute because I wasn't there, but he talked to my colleagues. And they reported, if I have it right, 10^{-10}. Well, 10^{-10} is the Q of some of the saturated absorption peaks. I don't know what the absorbing molecules were. (*Note added in proof*: The work referred to is by Bordé et al. at the Faculté des Sciences in Paris but the value 10^{-10} was not the Q, but rather the sensitivity of finding the center frequency).

Precision Wavelength Measurement of the Methane 3.39μm Saturated Absorption Line by Laser-Controlled Interferometry

R. L. Barger* and J. L. Hall**

Joint Institute for Laboratory Astrophysics of the University of Colorado and the National Bureau of Standards, Boulder, Colo. 80302

The methane 3.39 micron saturated absorption line has proven to be most accurately reproducible wavelength observed to date, with stability of 10^{-13} and reproducibility of better than 10^{-11}. However, the absolute wavelength of this transition has been known only to about 10^{-6}. Because this wavelength-reference system makes possible many new precision measurements, it has become interesting to measure its wavelength with an accuracy limited by that of the Kr standard. Also, recent advances point to the imminent measurement of the *frequency* of this line, so that a simultaneous knowledge of its wavelength will lead to the ideal measurement of the speed of light.

Since 3.39 μm is beyond the limit of photographic techniques, precision interferometry must incorporate photoelectric detection with an interferometer whose optical path is scanned on a precisely known scale.

We have developed a new, extremely accurate technique whereby an interferometer plate displacement by piezoelectric crystals is measured directly in terms of the beat frequency between a methane stabilized laser and one with a precisely controlled variable offset.

For the methane-krypton comparison, a plane-parallel Fabry-Perot variable length interferometer is being used. Parallel translation of the scanning plate is achieved with a three-point servo of cavity length to give simultaneous transmission fringes at each of the points for light from the offset laser. The frequency offset is then varied until transmission fringes are obtained successively for each of the krypton lines and methane-stabilized laser line. Thus, the cavity length in units of beat frequency is obtained for each fringe, and then the method of fractional fringes is used to calculate the wavelengths. Corrections for dispersion of phase shift on reflection are made by dealing with virtual spacers, obtained as the difference between pairs of physical spacers. Corrections are made for diffraction effects and lack of plate flatness by combining measurements made with three different apertures on physical spacers of different lengths to give the same virtual spacer.

In the first experiments, using coatings with finesse of about five to obtain reasonable transmission for the krypton light, we have obtained *accuracies* corresponding to about 10^{-3} *fringe*, with a *precision* for setting on individual fringes of nearly 10^{-5} *fringe*. The accuracy was limited by improveable parameters such as temperature stability. This technique can also be applied to long path interferometry for measurement of wavelength ratios of various stabilized laser lines to accuracies limited by stability of the lasers.

Our preliminary value for the vacuum wavelength of the $F_1^{(2)}$ component of the P(7) line in the ν_3 band of CH_4, referenced to the 6057 Å standard line of krypton, is

$$\lambda = 33922.3103 \text{ Å}$$

$$\sigma = 2947.9124 \text{ cm}^{-1}.$$

The apparent uncertainty in this wavelength value is 1×10^{-8}, although it must be emphasized that *this is a preliminary number.* [While all appropriate checks on systematic effects seem to be in order, it is perfectly possible that additional data and further checks on systematic effects will result in a correction significant compared with our apparent uncertainty of 10^{-8}. However, a correction as large as 3×10^{-8} seems unlikely.]

Key words: Methane; 3.39 micron saturated absorption line; ^{86}Kr.

---

* Staff member of Radio Physics Division, National Bureau of Standards, Boulder, Colo. 80302.
** Staff member of Laboratory Astrophysics Division, National Bureau of Standards, Boulder, Colo. 80302.

Interferometry for Wavelength Comparisons

K. D. Mielenz

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

For the calibration of Lamb-dip stabilized lasers with respect to the ^{86}Kr standard, an automated scanning Fabry-Perot étalon has been placed in operation which can be operated at lengths of 5 and 10 cm, uses a capacitance-bridge technique for length interpolation, and allows routine wavelength measurements with a precision of $2 \cdot 10^{-9}$

For the wavelength comparison of molecularly stabilized lasers to an expected precision of $1 \cdot 10^{-11}$, long Gaussian-beam cavities are illuminated with local-oscillator lasers whose frequencies are either servoed or measured with respect to the stabilized lasers themselves. This apparatus also permits step-scanning the ^{86}Kr-line profile.

Key words: Fabry-Perot interferometer; Gaussian-beam resonant cavities; krypton standard; stabilized lasers; wavelength measurement.

1. Introduction

In a series of measurements, conducted jointly by the national standards laboratories of Great Britain, the United States, and West Germany [1], it was found that different samples of Lamb-dip stabilized helium-neon lasers may emit wavelengths differing from one another by as much as 2 parts in 10^7. These discrepancies were attributed to differences in filling pressure, gas losses during prolonged operation, and similar causes. It was also demonstrated by these measurements that the wavelength of a particular laser of this type can be defined, by individual measurement, to within 5 parts in 10^9.

Hence it was concluded that an automated wavelength comparator should be developed, at NBS Washington, for the routine calibration of Lamb-dip stabilized lasers (and, if desired, conventional discharge-type interferometric sources) against the current primary standard of length. This instrument, which is described in section 2, below, has been completed and is now available for this purpose.

On the other hand, these discrepancies between Lamb-dip laser wavelengths have resulted in the development of improved lasers whose wavelength stabilities equal, and even surpass, that of the ^{86}Kr primary standard. Among them, those locked to an inverted Lamb-dip caused by saturated molecular absorption [2, 3, 4] offer promise of developing into a new network of internally consistent, highly refined wavelength standards. In response to this development we have established a second, more general interferometric facility for the pairwise comparison of stabilized lasers, and for the comparison of each of them with the ^{86}Kr standard. This apparatus, which is initially intended to compare methane and iodine stabilized lasers, is described in section 3.

2. Visible Wavelength Comparator

Figure 1 shows a block diagram of this apparatus. Its basic optical element is a further development of the scanning flat-plate Fabry-Perot étalon described earlier [5, 6]. The 5-cm fused-quartz interferometer plates are flat to $\lambda/100$ at 600 nm, and are aluminized over a 2.5 cm area to a reflectance of about 70 percent; this reflectance, as well as the plate separation (5 or 10 cm), being chosen in accordance with the optimum specifications for ^{86}Kr lamp illumination [7, 8]. The étalon is supported vertically inside an evacuated multiwalled housing incorporating water jackets and thermoelectric elements to maintain its temperature constant to better than 1/1000 °C during each measurement.

As the means to produce the scan motion over the required range of a few fringe orders, the étalon spacer incorporates a 5 cm section of piezoelectric ceramic tubing (PZT) having silver electrodes on both sides and expanding at a rate of approximately 500 V per order at $\lambda = 600$ nm. The outside electrode of this piezoelectric scanner is divided into three sections, so that three different voltages may be applied to adjust the parallelism of the étalon plates.

The nonlinearity, hysteresis, and time lag of the piezoelectric spacer element do not permit the d-c voltage applied to it to be used as a sufficiently precise measure of the scan motion. Therefore, a capacitance transducer is attached to the étalon to detect its motion, correct for the errors, and make possible linear interpolation of length. One capacitor plate is mounted on the outside of the top étalon plate and another is fixed inside the interferometer housing, so that a capacitance is formed which varies during the scan. This capacitance is compared with a fixed reference capacitor in a bridge, and the

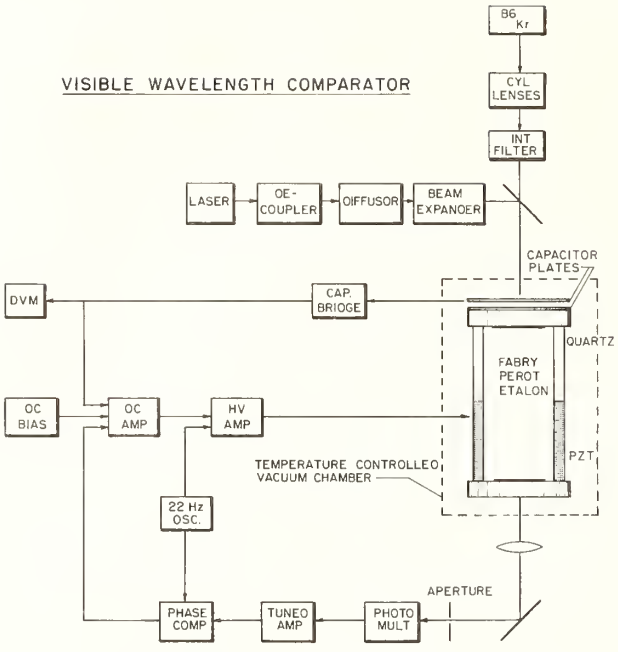


FIGURE 1. *Block diagram of visible-wavelength comparator.*

amplified bridge signal is then fed back degeneratively to the scanner in order to linearize its response and reduce the time lag.

The étalon spacer also includes a removable 5-cm section of quartz tubing, thus allowing the interferometer to be used at lengths of 5 and 10 cm to evaluate the dispersion of phase change on reflection of the aluminum plates. For the main intended use of the apparatus, the routine calibration of 633 nm laser wavelengths against the 606 nm ^{86}Kr line, this phase correction is quite small [9], so that the change of spacer lengths will have to be made only occasionally.

Light from the ^{86}Kr lamp is collimated by means of a cylindrical-lens system ensuring uniform interferometer illumination, and is then passed on to the étalon through multilayer dielectric filters isolating the 606 nm primary standard line, as well as the secondary-standard lines at 565, 665, and 670 nm which are used for order searching. The light from the laser is decoupled from the interferometer by a quarter-wave plate/polarization filter combination, is made spatially incoherent by passing it through a rotating ground-glass diffusor, and is then expanded to fill the interferometer aperture. On the output side of the étalon, a pinhole aperture is located in the image plane of a focusing lens, thus passing the innermost portion of the Haidinger ring pattern on to a photomultiplier tube. Pinholes of different size (0.5 to 1.5 mm diameter) are used to evaluate the correction due to nonuniformity of aperture illumination. The radiant flux from the 606 nm ^{86}Kr line received by the photomultiplier through a 1 mm pinhole is 4×10^7 photons per second, approximately.

To set on a given fringe, a predetermined dc bias voltage is applied to the piezoelectric scanner at the bridge-loop summing junction, thus producing an approximate setting. A 22 Hz ac voltage is superimposed to dither the interferometer length by roughly $\pm1/3$ order separation. The photomultiplier signal obtained is compared with the dither voltage by phase-sensitive detection, and the resulting derivative signal is added at high gain to the dc scan voltage in order to drive the interferometer to the fringe center. The corresponding capacitance-bridge signal is read out on a digital voltmeter as the final measure of fringe position. Figure 2 shows successive settings on a laser fringe from opposite directions, and indicates that such settings are possible within 1/5000 of an order separation, approximately.

For the measurement of an unknown wavelength, settings are made for successive-fringe pairs of each of the wavelengths illuminating the interferometer. The resulting readings from the digital voltmeter are processed by computer to yield the corresponding fringe fractions. The computer then performs an order search and, finally, computes the unknown wavelength. Repeated measurements of the same laser have resulted in individual wavelength values consistent with each other to ±2 parts in 10^9.

3. Interferometer System for Molecularly Stabilized Lasers

3.1. Design Principles

The high frequency stability of molecularly stabilized lasers indicates that meaningful measurements of their wavelengths may be carried out to at least 1 part in 10^{11}. Since conventional interferometric techniques are inadequate for such a high degree of accuracy, the interferometer system which we have designed and assembled for this purpose (fig. 3) incorporates a number of significant new design principles. These will now be explained, followed by a brief description of the apparatus itself.

a. Dual-Wavelength Operation

Since the two wavelengths to be compared will often be quite different from one another, the interferometer must be designed for simultaneous operation in widely separated wavelength regions. We have initially addressed the comparison of a methane-stabilized laser at 3.39 μm [3] with an iodine-stabilized laser at 633 nm [4]. This required the design of coatings having high reflection finesses at both wavelengths. The order of interference in the visible being five times greater than that in the infrared, it seemed desirable to make the visible finesse of these coatings five times smaller than their infrared finesse, so that the fringes of both wavelengths would become equally wide on a displacement scale. Purely metallic or dielectric coatings could not be designed to meet these requirements. Therefore, the interferometer plates used are provided with Au coatings overlayed with a dielectric stack of TiO_2 and SiO_2 layers one quarter wave thick at 3.39 μm. Their finesses are approximately 100 at 3.39 μm, and

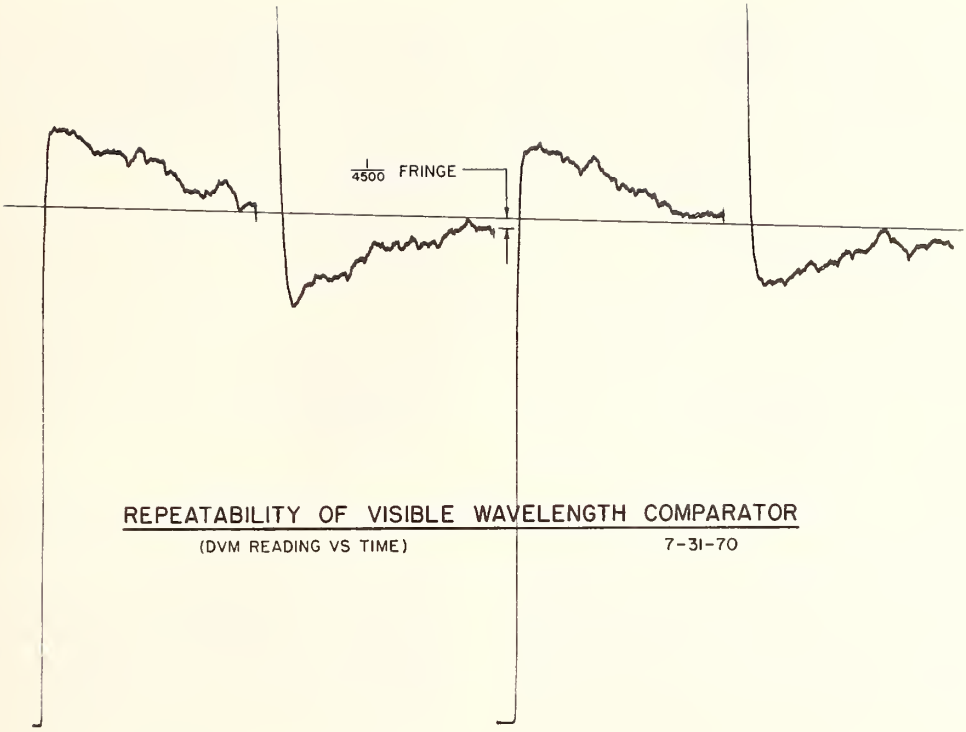


FIGURE 2. *Repeatability of fringe setting with visible-wavelength comparator.*

20 at 633 nm. Similarly, the apparatus includes a number of unusual beam splitters and reflectors (e.g., semitransparent at both wavelengths, or opaque at one wavelength and transparent at the other), whose design was not entirely trivial.

In order to facilitate infrared alignment, we have also included a red alignment laser whose beam is mode matched into the infrared cavities in a manner such as to simulate a "visible 3.39 μm beam." Using achromatic mode-matching optics, it is possible to employ the red light to make the initial adjustments for the infrared beam.

b. Gaussian-Beam Interferometers Used in Matched Pairs

In view of the infrared wavelengths involved, it is necessary to operate at path differences too large for flat-plate interferometers to be practical. Therefore, we employ curved-plate resonators, together with the appropriate mode-matching lens systems to excite their fundamental TEM_{00} modes. The necessary correction for reflection phase change is obtained using a second resonator having nearly identical mirrors, but operating at a different length. These two interferometers are used simultaneously, with an occasional interchange of their mirrors.

In order to avoid the complexity of separate mode-matching optics for the two interferometers they were designed as a "matched pair", utilizing the fact that the wavefront curvature of a Gaussian beam [10],

$$R(z)=z[1+(\pi w_0^2/\lambda z)^2], \tag{1}$$

is the same at two different distances z from the beam waist w_0, so that there are two mirror positions for which the resonator mode configurations are identical. In our particular case, each of the two interferometers has a 3.12 m concave mirror and a flat; their lengths are 0.62 and 2.50 m. They are excited through their flat end plates, using a common mode-matching system followed by a beam splitter.

The resonant wavelength of these interferometers is given by [10]

$$2L/\lambda=(q+1)+\beta/\pi, \tag{2a}$$

$$\beta=(2p+l+1)\cos^{-1}[(1-L/R)^{1/2}], \tag{2b}$$

where q, p, and l are the longitudinal and transverse mode orders, L is the interferometer length, and R is the concave-mirror radius. For a measurement of λ, the phase-shift term β/π must be determined in addition to the interferometer length L. This can be done by locking the interferometer to successive off-axial modes ($p, l\neq 0$), and then using eq (2b) to calculate β/π from the intermode beat frequencies thus obtained. For the two interferometers used, the expected values of β/π are, roughly 0.1 fringe ($L=$ 0.62 m) and 0.3 fringe ($L=2.5$ m), respectively.

c. Tuned Local Oscillators Servoed to Interferometers

For the high accuracies intended it is no longer possible to scan through successive fringe orders, and then rely on some auxiliary measurement of path length (such as by the capacitance method described in section 2, above). For this reason, the two inter-

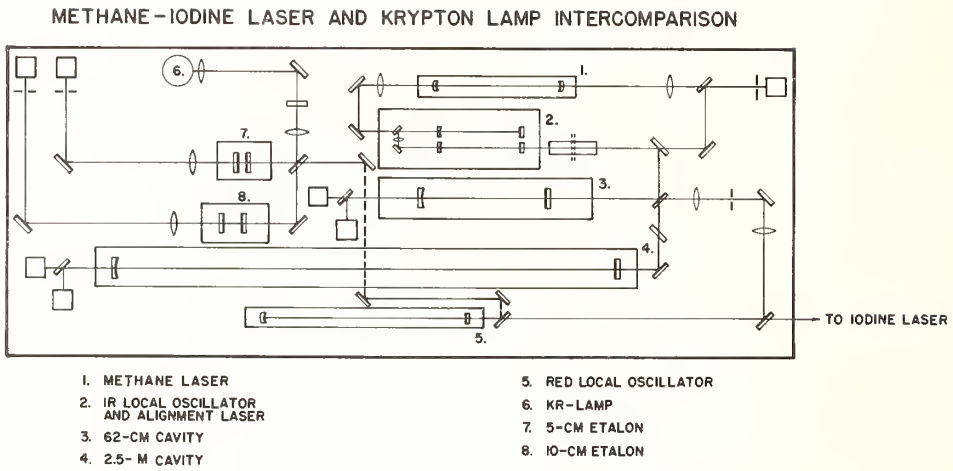


FIGURE 3. *Layout of apparatus for wavelength comparison of methane and iodine stabilized lasers and* ^{86}Kr *lamp.*

ferometers used are *not* illuminated by the two lasers to be intercompared, but by auxiliary lasers whose wavelengths can be adjusted such that the interferometers are resonant for both of them at the same time. In comparing the methane and iodine lasers, these are two helium-neon local-oscillator lasers operating at 3.39 μm and 633 nm, respectively; each of them servoed to yield maximum interferometer transmittance.

This procedure has the further important advantage of providing excellent decoupling of the two molecularly stabilized lasers from the two interferometers. It is well known that, even with considerable effort, there remains a small residue of frequency pulling in any laser-illuminated resonant-cavity interferometer. This is of too large a magnitude to be tolerable as a perturbation of the two molecularly stabilized lasers but is acceptable in the case of these local oscillators.

d. Integer Determination by Microwave Modulation

The two interferometers being illuminated by only two wavelengths, and these being widely separated, it would be difficult to determine their lengths by the customary methods of order searching. To overcome this difficulty, the local oscillator operating at 633 nm was designed in a manner so that it emits, in addition to its fundamental frequency ν, two sideband frequencies, $\nu \pm \omega$, produced by an intracavity KDP crystal modulated at a microwave frequency ω. Choice of this microwave frequency, such that the two sidebands are simultaneously resonant in the interferometer, then establishes the integral order number. This method is described elsewhere [11], and will not be discussed further in this paper.

e. Heterodyne Measurement of Excess Wavelengths

With the two local oscillators locked and the integral order numbers determined, the ratio of the two auxiliary wavelengths is known. The remaining task is to beat the 3.39 μm local oscillator with the methane laser, and the 633 nm local oscillator with the iodine laser. The two beat frequencies thus obtained play a role similar to that of fractional orders in conventional interferometry, and yield the final information necessary to determine the ratio of the methane and iodine wavelengths.

In realizing this concept, we take advantage of the high stability of the methane laser by establishing a frequency-offset lock between it and the infrared loeal oscillator, so that the interferometers and the red local oscillator also become controlled by the methane laser. This frequency-offset lock provides a convenient control capable of the high precision required, and greatly reduces the problems of interferometer instabilities, etc., that would be incurred otherwise.

3.2. Description of Apparatus

The layout of the interferometer system, as actually assembled, is shown in figure 3. All components are mounted on a 1.5×4 m vibration-isolation table. The two Gaussian-beam interferometers (3 and 4), mounted horizontally in evacuated stainless-steel tubes, are illuminated by the red and infrared local oscillators (5 and 2) by means of the mode-matching systems indicated. Also shown is the methane laser (1) and the optics required to beat its output with that of the infrared local oseillator (2). The latter is rigidly mounted together with a red alignment laser, whose output beam is mode-matched into the methane-laser and the infrared-local-oscillator cavities for the alignment purposes noted under 3.1a, above. The light from the red local oscillator (5) is mixed with that from the iodine laser (not shown) outside of the area in figure 3.

In view of the possibility that molecularly stabilized lasers may become future standards of length, it is necessary to connect their wavelengths with the ^{86}Kr scale, at least to the limit of accuracy of the present meter definition. For this purpose, the apparatus assembled also contains two scanning flat-plate Fabry-Perot interferometers (7 and 8),

operating at lengths of 5 and 10 cm, and illuminated by the red local oscillator (5) and a standard ^{86}Kr lamp (6). Again using the sidebands on the local-oscillator light to determine the order of interference, and interchanging mirrors to obtain the reflection phase correction, it is thus possible to measure the ^{86}Kr wavelength in terms of the local-oscillator wavelengths and, hence, in terms of the methane and iodine wavelengths. In addition, incrementing the microwave sideband frequency permits scanning the ^{86}Kr-line profile in order to detect its degree of asymmetry.

At the time of this Conference (August 3, 1970), all components of the apparatus described are on hand or immediately foreseen. Assembly of the system shown in figure 3 has begun.

4. Acknowledgments

The contents of this report are the results of an extensive group effort at NBS Washington. The present group leader is R. D. Deslattes, whose contributions to all aspects of this work are numerous. The electronic circuitry of the wavelength comparator described in section 2 was developed by H. D. Cook, aided by A. T. Funkhouser. W. G. Schweitzer and H. P. Layer, aided by D. K. L. Tan and J. L. Mikesell, have constructed the various lasers used with the interferometer system in section 3. K. F. Nefflen and C. Sidener have assisted in many optical and electronic problems. The interferometer coatings mentioned in section 3.1a were supplied by V. Costich of Spectra-Physics, Incorporated.

5. References

[1] Mielenz, K. D., Nefflen, K. F., Rowley, W. R. C., Wilson, D. C., and Engelhard, E., Appl. Optics **7**, 289 (1968).
[2] Lee, P. H., and Skolnick, M. L., Appl. Phys. Letters **10**, 303 (1967).
[3] Barger, R. L., and Hall, J. L., Phys. Rev. Letters **22**, 4 (1969).
[4] Hanes, G. R., and Dahlstrom, C. E., Appl. Phys. Letters **11**, 362 (1969).
[5] Mielenz, K. D., Stephens, R. B., and Nefflen, K. F., J. Res. Nat. Bur. Stand. (U.S.), 68C (Eng. and Instr.) No. 1, 1–6 (Jan.–Mar. 1964).
[6] Mielenz, K. D., Stephens, R. B., Gillilland, K. E., and Nefflen, K. F., J. Opt. Soc. Am. **56**, 156 (1966).
[7] Hanes, G., Can. J. Phys. **37**, 1283 (1959).
[8] Hill, R. M., Bruce, C. F., Aust. J. Phys. **15**, 194 (1962) and **16**, 282 (1963).
[9] Koester, C. J., J. Res. Nat. Bur. Stand. (U.S.), 64A (Phys. and Chem.) No. 3, 191–200 (May–June 1960).
[10] Kogelnik, H., and Li, T., Appl. Optics **5**, 1550 (1966).
[11] Bay, Z., this conference.

DISCUSSION

A. Javan: May I ask a question? You have mentioned that you have this marvelous system to compare the krypton standard with the helium neon standard. Is this one that you are volunteering to do for somebody who needs it badly? If I bring my laser would you measure the wavelength?

K. D. Mielenz: If you need it badly, yes. (*Laughter.*)

A. Javan: Yes? Wonderful. Very good. This is good to know. Thank you.

The Use of Microwave Modulation of Lasers for Length Measurements

Z. Bay

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

The diffraction phase shift, Φ, for the Gaussian TEM_{00} modes in a cavity formed by spherical mirrors of large aperture depends only on the cavity geometry. Thus the sidebands, $\nu \pm \omega$, of a microwave modulated laser propagate in an evacuated cavity with the same Φ, and the phase velocity of the beat wave (frequency 2ω) is c. This permits the establishment of a linear length scale based on the vacuum velocity of light. Tuning the cavity simultaneously to both sidebands makes its length an (easily known) integral multiple of half the beat wavelength, $\Lambda = c/2\omega$; but, the precision of the setting is given by the high cavity finesse for the optical λ. The setting of the interferometer length is periodic in the "small units" of $\lambda/2$. However, the λ scale is not linear, due to Φ. Comparison of these two scales permits the determination of the integral order number for $\lambda/2$ and also of Φ. The use of the microwave scale alone does not require the knowledge of λ. Thus, if (to eliminate the effects of phase shifts at reflection) two settings of the interferometer L_1 and L_2 are used, the difference of the two lengths is simply given by $L_2 - L_1 = n_2\Lambda_2/2 - n_1\Lambda_1/2$, even if a secular change in λ occurs between the two measurements. Due to its linearity, the use of the microwave scale appears preferable to that of the λ scale, if c is known with the precision of the present length standard. This can reasonably be expected in the near future. Since optical frequency comparisons and length measurements based on microwave frequencies avoid the use of a wavelength standard without any loss in accuracy, the adoption of a unified standard of time, frequency and length via a defined value of the velocity of light can strongly be advocated.

Key words: Linear length scale based on the speed of light; microwave modulation of lasers; tuning of optical cavities to microwaves.

1. Introduction

The use of high finesse interferometer cavities for precise length measurements, in which the cavity length is evaluated in terms of an optical wavelength, is made inconvenient by two requirements.

One is that a high optical order number, N, associated with the cavity length, L, has to be determined. For visible light and L of about 1 meter or larger, N is of the order of 10^6 or larger, making the application of fringe counting or optical multiplication methods inconvenient or difficult [1].

The other requirement is that the diffraction phase shift must be known. Its determination, if high accuracies are intended, requires additional efforts.

Both requirements are eliminated if instead of the direct laser light (wavelength λ, frequency ν) the sidebands of a microwave modulated laser (frequencies $\nu \pm \omega$, ω modulation frequency) are used in the length measurements. The length is measured in this case in terms of wavelengths of microwaves, while the precision and accuracy of the measurements is that associated with optical techniques.

2. Outline of Method

The diffraction phase shift, Φ, for self reproducing Laguerre-Gaussian waves in a cavity formed by spherical mirrors of large aperture depends only on the cavity geometry. Its expression (in units of order numbers) is

$$\Phi_{p,l} = \pi^{-1}(2p+l+1) \times \arccos\{[1-(L/R_1)][1-(L/R_2)]\}^{1/2} \quad (1)$$

where R_1 and R_2 are the radii of curvature of the mirrors, and p and l the radial and angular mode numbers of the off-axis modes [2]. (In this paper, use will be made of the axial TEM_{00} modes.)

Thus the sidebands, $\nu \pm \omega$, propagate in an evacuated cavity with the same Φ and the phase velocity of the beat wave (group velocity of the wave packets formed by the two waves) is c. This permits the establishment of a linear length scale based on the vacuum velocity of light.

The interferometer cavity is tuned to both sideband frequencies in the same way as in our experiments for the measurement of optical frequencies [3, 4]. Two servos are used. One tunes the length L of the interferometer (via the sidebands) to ν. The other servo adjusts ω to L according to the equation

$$2\omega = (n + \Delta\Phi_r)(c/2L) \quad (2)$$

where the integer n is the difference of the order numbers for the sidebands, and $\Delta\Phi_r$ is the difference of the reflection phase shifts for the sideband frequencies. In contrast to diffraction phase shifts, this difference is in general not zero. Expressing this difference as

$$\Delta\Phi_r = (\partial\Phi_r/\partial\nu)2\omega \quad (3)$$

with the derivative taken to be constant over the relatively small frequency interval of 2ω, eq (2) can be rewritten in the form

$$2L = n\Lambda + l_r \tag{4}$$

where $\Lambda = c/2\omega$ is the wavelength of the beatwave and $l_r = c\partial\Phi_r/\partial\nu$.

Thus, except for the small length, l_r, which is independent of ω, simultaneous tuning of the interferometer to the sidebands makes $2L$ an integral multiple of the beat wavelength. If, to eliminate l_r, two "telescoping" positions of the interferometer are used, then the difference of the two lengths (displacement of one of the mirrors) is given by

$$L_2 - L_1 = (n_2\Lambda_2/2) - (n_1\Lambda_1/2) \tag{5}$$

or

$$L_2 - L_1 = (n_2/2)(c/2\omega_2) - (n_1/2)(c/2\omega_1) \tag{6}$$

Thus by measuring ω_2 and ω_1 the length $L_2 - L_1$ is measured on a linear scale based on the vacuum velocity of light. The integers n_2 and n_1 are easily known if ω_2 and ω_1 correspond to centimeter or millimeter waves.

Note that no optieal λ appears in eqs (5) and (6). Thus, frequency stability of the laser is required only while each of the ω's is measured. Equations (5) and (6) remain valid even if secular changes in the laser frequency occur between the two measurements. The range of the frequencies required of the microwave equipment is rather limited. For example, a 5 percent variability of ω is sufficient to cover every $L_2 - L_1$, if $n_1 = 20$. For every $L_2 - L_1$ which is close to an integral multiple of $\Lambda_1/2$, ω_2 is nearly the same as ω_1.

The RF modulation of visible light for the measurement of distances via the velocity of light is well known and is widely used. The Geodimeter of E. Bergstrand, the Ultrasonic Light Modulator of A. Karolus and T. Helmberger, and the Mekometer of K. D. Froome and R. H. Bradsell [5] measure the phase difference between the outgoing and returning beatwave for the evaluation of the distance in units of the beat wavelength. The use of self reproducing standing waves in interferometer cavities provides for the exact repetition of phases here, makes the measurements exploitable with optical precision, and eliminates diffraction effects from the propagation of beatwaves.

3. Precision and Accuracy of the Measurements

The precision of the setting of ω is outlined in [3] and [4]. The operation of the ω servo can be made to be limited only by the shot noise of the photocurrents. Other noise, owing to short term fluctuations in ν and L, is strongly correlated in the signals corresponding to the two sidebands and drop out of the difference signal which is used to drive ω. Taking L between 1 and 2 meters, with a cavity finesse of about 500, photocurrents of about 10^{10} electrons per second, integration times of a few seconds, and ω of a few tens of GHz, the precision of ω can reach and exceed the one part in 10^{10} level. The precision increases with increasing length, integration time, and ω.

The ultimate accuracy of the measurements should be estimated from considerations related to departures of cavity parameters from exactness and their uncertainties. We treat briefly here irregularities of mirror surfaces, resulting in uncertainties in the length of the cavity. For simplicity we assume that other errors are made less than these with utmost care.

It appears that present day "superpolished" mirrors deviate from perfect geometric regularity in two respects [6, 7, 8].

(1) Across small distances of the order of an optical λ or less, the deviations can be characterized by randomly variable slopes of the order of unity. This lack of smoothness causes scattering of the light and diminishes the efficiency of the cavity transmission, but may be of little effect in changing the parameters of the cavity.

(2) There are slopes persistent over distances of millimeters or centimeters of the order $S \sim 10^{-6}$. These can alter the length of the cavity by a few nm when different spots of the same mirror are used in different alignments. They also may change the local radius of curvature for $R \sim 1$ meter by a part in 10^4 (which is also the limit of detection in a Foucault or Ronchi test).

The uncorrelated errors of a few nm for L_1 and L_2 would limit the accuracy of $L_2 - L_1$ in the telescoping displacement to about a part in 10^9 if $L_2 - L_1$ is a few meters. Possibly the accuracy can be increased by careful alignment of the beams to the same spot, but the deviations can not be completely eliminated from the measurements, because the Gaussian spot sizes are different for L_1 and L_2. It should be noted that this type of error occurs in the same way in length measurements based on an optical λ.

Another type of error has to be considered in measurements based on the sidebands. The Gaussian spot area of a self reproducing wave is proportional to λ. Thus the effective length of the cavity, averaged over the Gaussian amplitude distribution, is in general a function of λ, if mirror irregularities are present. For the two slightly different λ's of the sidebands, this can introduce a small phase shift which, according to eq (2), changes ω.

To estimate an upper limit for that phase shift we take the slope S across the Gaussian spot radius, w. The difference of the cavity lengths, as averaged over the two Gaussian spot areas is smaller than $Sw\Delta\lambda/\lambda = Sw\lambda/\Lambda$, where $\Delta\lambda/\lambda = \lambda/\Lambda$ is the relative difference of the spot areas. The quantity $Sw\lambda/\Lambda$, divided by the optical $\lambda/2$ is the phase shift which enters into eq (2). The relative change of ω for that phase shift is

$$\delta\omega/\omega = Sw/L. \tag{7}$$

For $w \sim 1$ mm, $S = 10^{-6}$ and $L = 1$ m eq (7) results in the error of one part in 10^9. The actual error is smaller than this, but it cannot be safely assumed that it is smaller than one part in 10^{10}.

It has to be noted that the same error occurs if

the length measurement is based on the optical λ, and Φ is measured via the off-axis modes of the interferometer. For different p and l, as seen from eq (1), the diffraction phase shifts differ by multiples of Φ_{00} and those phase differences can be measured by shifts in the eigenfrequencies of the cavity.

However, the Laguerre-Gaussian wave forms of the off-axis modes occupy spots on the mirrors which for increasing p and l extend increasingly farther away from the center of the axial mode. For the lowest mode numbers, p and l, the average displacement from the center is of the order of w. Thus, the relative change of the cavity length in such measurements is also of the order of Sw/L.

Therefore we conclude that length measurements, either based on an optical λ, or based on microwave modulation of lasers, are of similar accuracies. It is hard to expect that the accuracy of a single measurement can reach the one part in 10^{10} level for lengths of a few meters. Since the errors considered are of statistical character, averaging of many measurements in different alignments may reach or pass that limit. A break-through in mirror polishing techniques would help in the same way in both methods.

Another interesting conclusion can be drawn from this for our optical frequency measurements [3, 4], using microwave modulation techniques. Those deal simultaneously with two scales, one based on the optical frequency ν which also includes Φ, the other based on microwave frequencies. Thus the accuracy of those frequency measurements is similar to the accuracies with which an optical λ, however well defined, can be utilized in length measurements or in the comparison of two different λ's, however well both are defined.

It is interesting to note that the simultaneous appearance of the two scales in our experiments leads to an easy determination of the optical order numbers, N, and it allows one to pursue the frequency measurements to somewhat higher accuracies than those achieved in measurements based on optical wavelengths.

The setting of the interferometer is periodic in the "small units" of $\lambda/2$. However, the λ scale is not linear, since Φ is a function of length. Expressing the same length $L_2-L_1=L_{12}$ on both scales, we obtain

$$(N_2-N_1+\Phi_2-\Phi_1)\lambda=n_2\Lambda_2-n_1\Lambda_1=2L_{12}. \quad (8)$$

Since c is known at present to about a part in 10^6, L_{12} can be calculated from the measurements of ω_1 and ω_2 to that same accuracy. Suppose that λ is also known to that accuracy, or better. From the approximate knowledge of the cavity parameters $\Phi_2-\Phi_1$ can be easily known to better than one percent. Thus if in the first application of eq (8), for a short L_{12}, N_2-N_1 is less than 10^5, that number can be determined to one unit. Then N_2-N_1 can be multiplied up by the ratios of displacements (from which c drops out) in successive steps, in which new estimates of the differences in Φ are used.

This method of determination of optical order numbers via modulation techniques can successfully replace fringe counting and optical multiplication methods for long interferometers, especially when no neighboring laser lines are available for the use of the method of "exact fractions" (for which the sidebands serve here). This can be useful in λ-comparison experiments.

After the order numbers, N, are known, eq (8) can now be applied for the determination of the diffraction phase shifts, and of ν, in terms of the microwave frequencies. Accepting only the functional dependence of Φ_{00} on the cavity parameters, measurements of ω at a proper number of points in the telescoping procedure allow the computation of the cavity parameters, and that of Φ and ν. The quantity ν results from such computations as the best statistical fit. Note that no off-axis modes are used in those measurements.

For short lengths of much below 1 meter, the microwave modulation techniques may not compete in simplicity with length measurements based on an optical λ. Above 1 meter, and up to the limit of lengths at which interferometer cavities are workable, the microwave techniques appear preferable.

Of course, length measurements via microwave modulation of lasers, presuppose the knowledge of c. It can be expected that c will be known in the near future to an accuracy limited by the present length standard.

Moreover it seems that, since length measurements based on microwave frequencies and the vacuum velocity of light avoid the use of a new wavelength standard without any loss in accuracy, they represent a strong point for the acceptance of a unified standardization of time, frequency and length via a defined value of c.

4. Problems of Standardization

After describing an experimental scheme for the measurement of optical frequencies, relating them to microwave frequencies, we pointed out [3, 9, 10] that the method can be used to establish a unified time-length scale via c, based on a single resonance transition which defines the unit of time. It is obvious that the knowledge of the frequency in terms of the standard second and simultaneous measurement of the corresponding wavelength in terms of the length standard results in refining the value of the velocity of light in meters/second.

Since it is expected that the accuracy of the optical frequency measurements will surpass that of the present length standard, the accuracy of c will be limited to that of the present meter. A definition of c (compatible with the present meter, but otherwise arbitrary) would result in a new definition of the meter. Through the knowledge of their frequencies, laser reference lines for spectroscopy and for precise length measurements would be available in any part of the optical spectrum without the need to define one particular wavelength as a new length standard. Also, time of flight measurements (in terrestrial and space radar) could be translated directly into distances via the definition of c.

The use of a fundamental constant of nature to

reduce the number of independently defined units of measurements certainly has an aesthetic appeal [11, 12, 13]. On the other hand, the acceptability of any proposed scheme has to be judged on the basis of operational considerations.

The interferometric comparison of λ's and optical frequency measurements using interferometers and microwave modulation techniques are limited in accuracy (by interferometer imperfections) in a similar way. Therefore, this may strongly suggest the adoption of a unified standard for time, frequency, and length via a defined value of c. That system of standardization can be advocated as soon as c is measured to the accuracy of the present length standard with respect to its definition.

It should be noted that if direct frequency measurements, based on phase locked frequency multiplier chains between optical and microwave frequencies materialize throughout the visible spectrum, they will provide for frequency comparisons with higher accuracies than λ comparisons or our frequency measurements, using interferometers. They will not directly help length measurements, the accuracy of which still will be dependent on cavity imperfections.

5. Conclusions

(1) The use of the sidebands of microwave modulated laser light in interferometric length measurements permits the establishment of a linear length scale based on the vacuum velocity of light. The length is conveniently expressed in terms of microwave frequencies and the measurements are free of diffraction effects.

(2) Comparison of the optical wavelength scale and the linear scale based on c permits the determination of optical order numbers and of diffraction phase shifts in interferometers.

(3) The limiting accuracies, as dependent upon interferometer imperfections, are similar in measurements based on optical wavelengths and in measurements based on optical frequency determinations via microwave modulation techniques.

(4) Since optical frequency comparisons and length measurements based on microwave frequencies avoid the use of a wavelength standard without any loss in accuracy, the adoption of a unified standard of time frequency and length via a defined value of the velocity of light can be advocated as soon as c is known to the accuracy of the present length standard.

6. References

[1] Baird, K. M., Metrologia **4**, No. 3, 135 (1968).
[2] Kogelnik, H., and Li, T., Proc. IEEE **54**, No. 10, 1312 (1966).
[3] Bay, Z., and Luther, G. G., Appl. Phys. Letters **13**, No. 9, 303 (1968).
[4] Bay, Z., and Luther, G. G., these Proceedings.
[5] Reviewed by Froome, K. D., and Essen, L., The Velocity of Light and Radio Waves (Acad. Press London and New York, 1969).
[6] Tolansky, S., Surface Microphotography (New York Interscience Publ. 1960).
[7] Tolansky, S., Microstructure of Surfaces (Amer. Elsevier Publ. Co., New York, 1968).
[8] Twyman, F., Prism and Lens Making (Hilger and Watts Ltd., London; University Press, Glasgow, 1957).
[9] Bay, Z., Internal NBS Report, 1965, (unpublished).
[10] Bay, Z., and Luther, G. G., NBS Technical Note 449, 30 (1968).
[11] Planck, M., Theorie der Wärmestrahlung, p. 163, Barth, J. A., Leipzig, 1906.
[12] Townes, C. H., Advances in Quantum Electronics, p. 3, J. R. Singer, (New York, Columbia University Press, 1961).
[13] Costa de Beauregard, O., Metrologia **4**, No. 3, 144 (1968).

DISCUSSION

A. Javan: There is one measurement, the speed of light, in which one can say that there are absolutely no duplications of effort. If two different groups attempt it or one had different approaches, it would be useful to check against each other. Well, that's the speed of light. So it's marvelous to have a different way of doing these measurements, and probably there will be other measurements that could be done.

Z. Bay: Yes. That is what is done now in the Bureau. We are doing lambda comparisons in one group, and we are doing our frequency measurements.

A. Javan: That's right. All these independent techniques eventually will serve to be checks on each other.

The Measuring of Optical Frequencies and the Velocity of Light

Z. Bay and G. G. Luther

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Experiments for the electrooptic modulation of a 632.8 nm He–Ne laser at X-band microwave frequencies, and for the simultaneous tuning of a Fabry-Perot cavity to both optical sidebands of the modulated laser have been reported. The optical frequency, ν, is thus related to the microwave frequency, ω. In recent developments, the microwave frequency is derived from the NBS "in-house" frequency standards. A sweep of ω, needed to establish the precise relationship between ω and the length of the cavity, is produced by phase modulation of the reference, such that the number of cycles in ω is conserved and the average ω is precisely known. To determine simultaneously ω and λ, the microwave modulated laser is locked with a known frequency offset to a Lamb-dip stabilized 632.8 nm laser the wavelength of which is measured in terms of the ^{86}Kr standard. In preliminary experiments, with the X-band frequency, $\omega \sim 10$ GHz, the finesse of the Fabry-Perot of about 600, length about 20 cm, photoelectric currents of a few times 10^7 electrons per second, and using averaging times of about a second, the precision in ω of better than a part in 10^8 could readily be achieved. This would permit the measurement of ν and c to the same accuracy with the reasonable expectation that those accuracies will exceed the precision of the length standard in future experiments. Possible applications for standardization problems will be discussed.

Key words: Frequency measurement of laser; intra-cavity modulation; velocity of light measurement.

1. Introduction

This is a report on an experiment now underway at NBS, Washington, to measure optical frequencies and the speed of light by means of a technique whieh utilizes a modulation of a laser beam at microwave frequencies. Some implications to problems in standardization are also discussed.

We have shown [1, 2] that the frequency of any laser oscillation can be related to the time standard without the use of a frequency multiplying chain to connect the optical to microwave frequencies. The connection is brought about by microwave modulation of the laser. The method can be used either: (a) to stabilize the laser frequency, or (b) to measure the laser frequency, stabilized otherwise, in terms of the microwave frequency.

2. Outline of Method

By modulation of the laser light (frequency ν) at the microwave frequency (ω) the sidebands $\nu \pm \omega$ are generated.

The two sidebands then pass through a Fabry-Perot cavity and to a detector. The length of the Fabry-Perot cavity (L) is adjusted so that the interferometer is simultaneously resonant to both of the sidebands. When this condition is met, then

$$\nu + \omega = N_+ (c/2L)$$
$$\nu - \omega = N_- (c/2L) \qquad (1)$$

and

$$\nu = [(N_+ + N_-)/(N_+ - N_-)]\omega \qquad (2)$$

where N_+ and N_- are the order numbers of the two sidebands in the Fabry-Perot and c is the speed of light. Diffraction and reflection phase shifts are ignored in these equations, but will be discussed in section 4.

Hence the optical frequency is determined by measuring a microwave frequency which is directly relatable to the frequency standard, and by determining the two integers, N_+ and N_-, and the corrections to them. The fluctuations in the (servoed) value of ω are a measure of the experimental precision. The extent to which the corrections to the N's are known is a measure of the accuracy of the connection between ν and ω. Note that the measurement of ν in terms of ω does not require knowledge of the velocity of light or the length of the cavity.

The exact resonance condition is established by two servo loops using the light of the modulated laser as shown in figure 1. The length of the Fabry-Perot cavity is modulated at a frequency γ_L (approximately 40 kHz in experiments in progress). The average length of the Fabry-Perot cavity is changed to minimize the first harmonic of the γ_L signal. Simultaneously, the microwave frequency, ω, is phase modulated at a frequency γ_ω (near 2 kHz in experiments) and ω is changed to minimize the first harmonic of γ_ω. The two servo loops make the Fabry-Perot resonant at both optical sidebands.

It is instructive to rewrite eq (2) in the form

$$\nu = (N/n)2\omega \qquad (3)$$

where $N = (N_+ + N_-)/2$ is the order number of ν, and $n = N_+ - N_-$ is the order number belonging to the

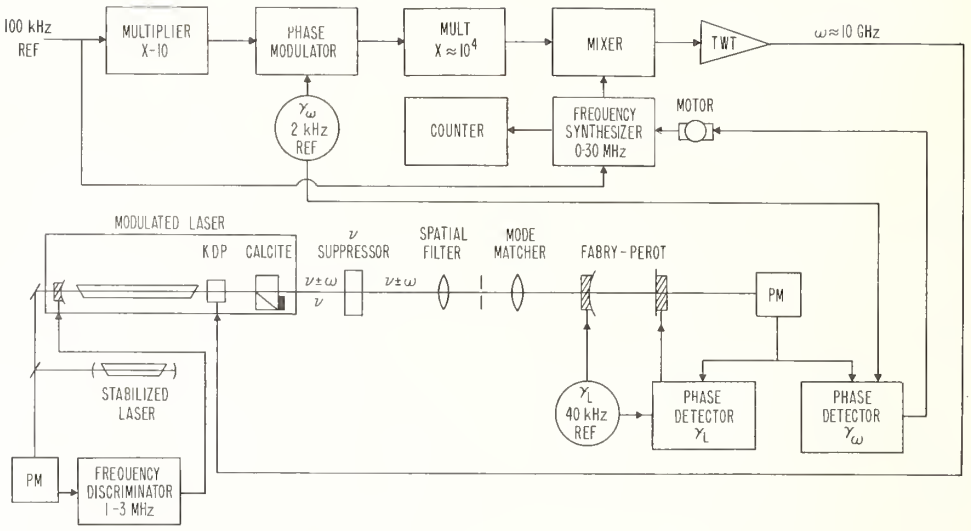


FIGURE 1. *Block diagram of the system used to measure the frequency of a stabilized laser.*
The system can also be used to lock the laser to the frequency standard.

beat frequency between the sidebands, 2ω. Thus, the interferometer appears to be tuned simultaneously to the optical and to the microwave frequency, even though the unmodulated laser beam, at frequency ν, is not used in the experiment. (In practice, for best signal-to-noise ratio, N is made a half integer to prevent any admixture of ν from passing the interferometer.)

The precision of locking L to ν (or vice versa) is similar in this experiment to that in other experiments in which a Fabry-Perot is locked to a laser. The accepted rule-of-thumb is that a hundredth (under quiet conditions a thousandth) of the width of the transparency curve of the interferometer may be taken for the limiting precision. For a finesse of a few hundred, and for $N \sim 10^6$ this gives a precision for L or for ν of the order of a part in 10^{10}, or possibly 10^{11}.

Off-hand it might be thought that the precision of setting ω to L is far inferior to that of setting L to ν, because of the relative smallness of ω as compared to ν (in our experiments $\nu \sim 5 \times 10^{14}$ Hz for the 632.8 nm He-Ne line, while $\omega \sim 10^{10}$ Hz is used for the modulation frequency).

However, there are favorable circumstances which can be utilized to achieve high precision in the setting of ω: (i) Since the sidebands are generated by microwave modulation of ν, their separation (2ω) on the frequency scale has the stability of the time standard; (ii) By the modulation scheme, applied to the phase of ω at the frequency γ_ω, the two frequencies, $\nu \pm \omega$, move by the same amount in opposite directions. Thus the two first harmonic amplitudes belonging to the sidebands are strongly correlated. Their difference, which is picked up in the phase detection at γ_ω and which is used to set ω, is insensitive to fluctuations in ν, in L, or in the airpath and optics (used for mode-matching) between the laser and the interferometer. Therefore this signal can be "smooth," down to the shot noise limit of the photocurrent; (iii) to diminish that noise, long time constants can be used in the servoing of ω (one second or greater). This has no detrimental effect on the operation of the whole system, since L is continuously servoed to ν.

With photoelectron currents of 10^{10} electrons per second in the sidebands and a one second integration time the relative shot noise fluctuation is 10^{-5}. Thus the precision of the setting of ω is 10^{-5} of the frequency bandwidth of the interferometer, or a few Hz for $L \sim 2$ meters and the finesse $F = 600$. The relative precision in the setting of ω is then a few parts in 10^{10} for $\omega = 10^{10}$ Hz. An increase by one order of magnitude can be achieved by increasing ω to 10^{11} Hz. It is well known that KDP crystals perform equally well at even higher frequencies. By a combination of further increases in L, in the averaging time, and in the photocurrents, the precision in the setting of ω may come close to that of the time standard.

Problems of accuracy, similar to those encountered in other uses of interferometers, will be dealt with later.

To obtain the speed of light it is necessary to measure simultaneously the frequency and wavelength of the laser output. This is accomplished by locking the laser frequency ν, with a predetermined difference frequency, to a stabilized He–Ne laser. The wavelength of the latter can be determined relative to ^{86}Kr by techniques such as reported by K. Mielenz at this Conference [3].

3. Experimental Procedure

The equipment and operation of the system, as schematically outlined in figure 1, is as follows:

A laser containing an intra-cavity modulator is locked with some frequency difference (1–3 MHz)

to a stabilized laser (at present, a Lamb-dip stabilized laser is used but in principle any stabilized laser, e.g., an iodine or methane stabilized one may be used), whose wavelength can be compared to the ^{86}Kr standard. The laser is amplitude modulated at approximately 10 GHz via intra-cavity modulation. The KDP crystal, when the field is applied, converts some of the light in the cavity to the orthogonal polarization, which is coupled out by a plane parallel calcite piece, which has a totally reflecting mirror cemented to its rear surface. Since the light coupled out by the calcite is orthogonal in polarization to that of the laser, no additional decoupling scheme is used.

An area of the mirror is clear to allow the modulated light to leave the cavity. This light then passes through a short plane parallel Fabry-Perot in order to suppress the unmodulated light and to equalize the intensities of the two sidebands. This short Fabry-Perot is made from a plane parallel quartz plate 5 mm thick and dielectric coated on both surfaces to 70 percent reflectivity. This gives a free spectral range of 20 GHz (2ω) and a finesse of something greater than 8, with an efficiency of nearly 100 percent. A spatial filter, consisting of a 20 power microscope objective and a 10 μm pinhole, is used to smooth out the wave front. A 25 mm focal-length lens is used to mode match this 10 μm pinhole into the Fabry-Perot interferometer. The output of the interferometer is detected by a photomultiplier. The interferometer is locked to the modulated laser sidebands by modulating its length at γ_L (40 kHz) and phase detecting the output of the photomultiplier using this γ_L as a reference frequency. The length of the Fabry-Perot is servoed in order to make this signal zero. At the same time, the microwave frequency is phase-modulated at γ_ω with the resulting frequency deviation chosen to be approximately equal to the width (at half maximum) of the transparancy function of the Fabry-Perot. In the modulation of ω, phase modulation is chosen because this modulation leaves the average frequency unchanged even though the modulation is not symmetric.

The phase modulation is done at 1 MHz, which is derived from a 100 kHz reference, then multiplied up to 10 GHz and mixed with the output of a frequency synthesizer. The range of the frequency synthesizer is 0 to 30 MHz. The mixed signal is then amplified to about 1/2 W via a traveling wave tube amplifier and then coupled into the KDP. The output of the photomultiplier phase detected at γ_ω is used to servo the frequency of the synthesizer so that the signal at γ_ω is zero. The 0 to 30 MHz output frequency of the synthesizer is determined by a frequency count. These frequency measurements constitute the information output of the system. At the present time the standard deviation in the frequency, for measurement periods of 1 second duration, is 70 Hz, which is 7 parts in 10^9 of ω. This is about the shot noise limit for 10^8 photoelectrons/second, and is about 10^{-4} of the band width of a 20 cm Fabry-Perot with a finesse of 600. This present result with a short Fabry-Perot amounts to approximately the accuracy to which the ^{86}Kr wavelength can be used. No doubt this precision can be increased by increasing the length of the interferometer, the integration time, the modulation frequency, or all of these.

4. The Accuracy of the Frequency Measurements

The diffraction phase shift, Φ, and the reflection phase shift of the mirrors, Φ_r, were ignored in eq (3), which should be modified to be

$$\nu=[(N+\Phi+\Phi_r)/(n+\Delta\Phi_r)]2\omega \qquad (4)$$

where $\Delta\Phi_r$ is the difference of the reflection phase shifts for the two sideband frequencies. The diffraction phase shifts for the self reproducing TEM_{00} modes of a cavity formed by spherical mirrors of large aperture are dependent only on the cavity geometry [4].

In order to eliminate Φ_r and $\Delta\Phi_r$, two settings of the cavity at L_1 and L_2 are used and the corresponding ω_1 and ω_2 determined. Then

$$\nu=\frac{N_2-N_1+\Phi_2-\Phi_1}{(n_2/2\omega_2)-(n_1/2\omega_1)}. \qquad (5)$$

The question of accuracy of ν, as obtained from eq (5), is treated in another report to this conference [5]. The numerator of eq (5), multiplied by λ, represents the length $2(L_2-L_1)$ as measured in terms of the optical wavelength. The denominator multiplied by c, gives the same length in terms of the microwave frequencies and the vacuum velocity of light. The accuracies of both expressions are estimated in reference [5] by taking into account surface irregularities of present "superpolished" mirrors, under the simplifying assumption that other errors may be eliminated, albeit with difficulty.

Bay [5] indicates that the accuracies of length measurements via an optical λ, and via microwave frequencies and c, are of similar order.

It then follows from eq (5) that the limiting accuracy of our frequency measurements is similar to the accuracies with which an optical λ, however well defined, can be utilized in length measurements or in the comparison of two different λ's, however well both are defined.

It should be noted that future improvements in the interferometric techniques such as a break-through in mirror polishing, or the use of very long interferometers, or statistical averaging of many measurements, help in λ-comparison and in frequency-comparison experiments in the same way. Although these improvements will help in the comparison of optical frequencies they cannot help in the redetermination of c since the expected accuracy of the frequency measurements even now will exceed the accuracy of the present length standard with respect to its definition.

Therefore, a unified standard for frequency and lengths via a defined value of c is recommended.

5. Conclusions

(1) The limiting accuracies, due to interferometer imperfections, are similar in λ comparisons, length measurements, and in the optical frequency measurements via microwave modulation techniques.

(2) The precision of our optical frequency measurements can be made high enough to utilize these limiting accuracies.

(3) On the basis of the foregoing conclusions the adoption of a unified standard for time interval and length via a defined value of the velocity of light is strongly advocated.

6. Acknowledgments

We are grateful to R. H. Carpenter, R. O. Stone, and K. W. Yee for assistance in the development of the electronic systems.

7. References

[1] Bay, Z., and Luther, G. G., Bull. APS. Ser. II, **14,** No. 4, 620 (1969).

[2] Bay, Z., and Luther, G. G., Appl. Phys. Letter **13,** No. 9, 303 (1968).

[3] Mielenz, K., Interferometry for Wavelength Comparisons, these Proceedings.

[4] Kogelnik, H., and Li, T., Proc. IEEE **54,** No. 10, 1312 (1966).

[5] Bay, Z., The use of Microwave Modulation of Lasers for Length Measurements, these Proceedings.

Defining the Speed of Light: A Combination Time, Frequency, and Length Standard: Recent Progress Toward Measuring the Frequency of Visible Light

K. M. Evenson, J. S. Wells, and L. M. Matarrese

National Bureau of Standards, Radio Standards Physics Division, Boulder, Colo. 80302

The frequencies of the water vapor laser at 3.8 and 10.7 THz (78 μm and 28 μm) and of the P(18) and P(20), 28 THz (10.6 μm) lines of the CO_2 laser have been measured in this laboratory. This was done by generating a beat note between the unknown radiation and combinations of various harmonics from lower frequency laser and klystron radiations impinging on a tungsten catwhisker-on-nickel diode. Efforts are presently underway to measure the frequency and wavelength of the methane-stabilized 88 THz (3.39 μm) He–Ne laser. Current estimations are that the value of c derived from this combined measurement will be better than 1 part in 10^8. It would then be possible to define c as exactly this number, and use the present time and frequency standard as a length standard also.

Key words: Laser; length standard; velocity of light.

Recent progress in the measurement of laser frequencies [1, 2] and in stabilizing lasers to saturated absorptions [3, 4, 5] suggest the possibility of using a single standard for time, frequency, and wavelength [6]. The present length, time and frequency standards are defined as:

$1m \equiv 1,650,763.73$ wavelengths of the $2_{p10}-5d_s$ transition of ^{86}Kr.

$$\begin{pmatrix} \lambda_{st} = 605.780211 \text{ nm} \\ \nu_{\lambda\ st}{}^* = 494.8866 \text{ THz} \end{pmatrix}$$

1 second $\equiv 9,192,631,770$ oscillations of the $F=4$, $m_F=0$ to $F=3$, $m_F=0$ transition of the fundamental state $^2S_{1/2}$ of ^{133}Cs

$$\begin{pmatrix} \nu_{st} = 0.009,192,631,770 \text{ THz} \\ \lambda_{\nu\ st}{}^* = 32,612,260 \text{ nm.} \end{pmatrix}$$

Wavelength comparisons can generally be made in the visible to a few parts in 10^{10}; however, at about 10 μm, the diffraction and phase shift corrections limit visible to infrared wavelength comparisons to about a part in 10^8. On the other hand, frequency comparisons have limits imposed only by the coherence of the sources and noise limits in the harmonic generator mixers. (These limits seem to be smaller than a part in 10^{11}).

Current measurements [3, 4] on the methane stabilized He–Ne laser at 3.39 μm yield an accuracy of 1 part in 10^{11} and a precision of better than 1 part in 10^{13}; the accuracy [5] of the iodine stabilized 632.8 nm He–Ne laser is better than 2 parts in 10^9. Both of the devices are not only more coherent, but more accurate than the present wavelength standards. Thus, one can immediately switch to either one of these devices and have a better length standard; or, if the frequency of either of these devices can be measured directly, one could obtain an extremely accurate value of c from $\nu\lambda$; c could be defined as exactly this value and then one could use the most stable source, (whether it be the present methane stabilized He–Ne laser, the cesium beam,

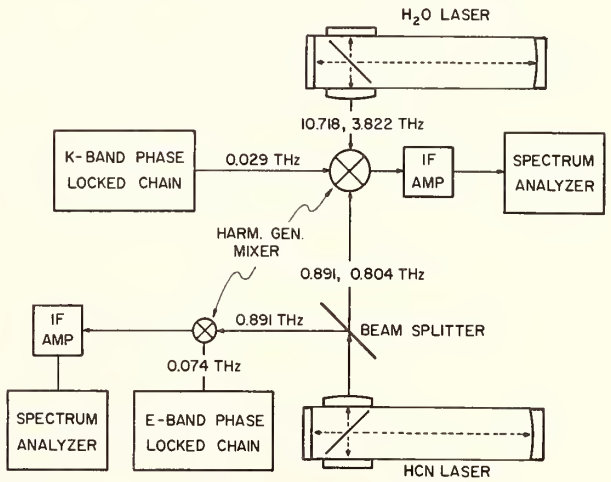


FIGURE 1. *Diagram of 28 μm and 78 μm frequency measuring experiment.*

* Using $c = 2.9979250(10) \times 10^8 m \cdot s^{-1}$.

TABLE 1. *Summary of laser frequency measurements*

Frequency ν_x (THz)	Wavelength λ_x (μm)	Power available (mW)	Type of laser	Laser			Laser	Klystron		Ref.
				n	ν_1 (THz)	m	ν_2 (THz)	l	ν_3 (THz)	
0.0106[a]		200								
.0742		100						7	0.0106	
.8907606	337	100	HCN					12	.0742	7, 1
.80475	373		HCN	1	0.891			3	.029	7, 1
3.821775	78	20	H_2O	6	.891	−2	0.805	3	.029	1
10.718073	28	350	H_2O	12	.891			1	.029	1
28.306251	10.6	2000	CO_2	3	10.718	−1	3.821	−1	.027	2
28.359800	10.6	2000	CO_2	3	10.718	−1	3.821	1	.026	2
32.176084	9.3	pulsed	CO_2	3	10.718			1	.022	8
32.134269	9.3	pulsed	CO_2	3	10.718			−1	.020	8
56	5	pulsed	CO	2	26					9
88.37637[b]	3.39	50	He–Ne[a]	8	10.718	3	.891	−1	.040	In progress

[a] X-band klystron—measured in counter.
[b] From wavelength measurements.

or whatever) as a combination length, frequency, and time standard.

It would be conceptually possible to compare the frequency of this primary standard of length, frequency, and time to the frequency of a secondary standard, and thusly, achieve a secondary standard of length, time, and frequency. In practice the only difference between defining λ or defining c is whether one prefers to have only 9 digits (plus zero) in the value of c or in the definition of length. A single standard, of course, would require frequency synthesis up to the visible region of the electromagnetic spectrum, and we would like now to summarize recent progress aimed at this goal.

Since the first laser frequency measurement at 0.89 THz early in 1967, the upper limit to which frequencies have been measured has expanded rapidly to a present value of 55 THz (5 μm).

In order to measure an unknown laser frequency, ν_x, one must add harmonics (n and m) from laser lines with frequencies ν_1 and ν_2, plus harmonics of a klystron to achieve a frequency coincidence; thus:

$$\nu_x = n\nu_1 \pm m\nu_2 \pm l\nu_3.$$

A summary of many of the laser frequencies presently measured, and the one presently in progress is shown in table 1. The work in this laboratory has concentrated on cw lasers, while the work at MIT under the direction of A. Javan has used mainly pulsed lasers.

A block diagram of the experimental arrangement used to measure the water vapor laser frequencies is shown in figure 1. A single diode acts as a combination harmonic generator and mixer for the entire combination of frequencies. A conventional tungsten catwhisker on silicon diode was used up to about 2 THz, while a tungsten on nickel diode was used at higher frequencies (this diode works well at the lower frequencies also). The success of the cw measurements was made possible through the use of three major improvements in increasing the currents on the diode: (1) the use of large lasers (8 m long), (2) variable coupling Michelson lasers for the HCN and H_2O lasers as shown in figure 2, and (3) improved coupling to the catwhisker with the use of long antenna theory [10]. General criterion for the characteristics of a suitable harmonic generator and mixer diode junction have now been established.

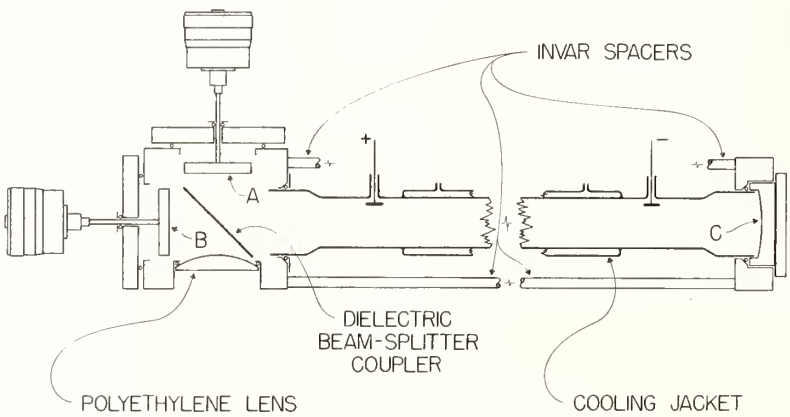


FIGURE 2. *Variable coupling Michelson far infrared laser.*

Currently an attempt is being made to measure the frequency of the He–Ne laser as shown in the last line of the table. This particular scheme is extremely advantageous since only two lasers are used to complete the chain of frequency synthesis to 88 THz. Radiation from all of the various sources is now coupled to the whisker diode, and quite acceptable rectified signals have been obtained which gives encouragement for the eventual success of this measurement. After the free-running-laser's frequency has been measured, a methane stabilized laser will be substituted. A precise measurement of its frequency coupled with current wavelength measurements by R. Barger and J. Hall will yield an extremely accurate value of the speed of light. The initial frequency measurements should be accurate to a few parts in 10^9; meanwhile progress is also underway to stabilize all of the lasers in this chain to yield an even more accurate frequency measurement.

At 28 THz and above, bulk optical second harmonic generation has been used, so that it is quite apparent, that even if the diode is eventually limited in response, that this other method will eventually lead to the measurement of frequencies in the visible.

References

[1] Evenson, K. M., Wells, J. S., Matarrese, L. M., and Elwell, L. B., Appl. Phys. Letters **16**, 159 (1970).
[2] Evenson, K. M., Wells, J. S., and Matarrese, L. M., Appl. Phys. Letters **16**, 251 (1970).
[3] Barger, R. L., and Hall, J. L., Phys. Rev. Letters **22**, 4 (1969).
[4] Barger, R. L., and Hall, J. L., Proceedings of the 23rd Annual Symposium on Frequency Control, Fort Monmouth, New Jersey, May 1969, p. 306.
[5] Haines, G. R., Baird, K. M., Metrologia **5**, 32 (1969).
[6] Simkin, G. S., Measurement Techniques **10**, 1308, (1970).
[7] Hocker, L. O., and Javan, A., Phys. Letters **25A**, 489 (1967).
[8] Danen, V., Sokoloff, D., Sanchez, A., and Javan, A., Appl. Phys. Letters **15**, 398 (1970).
[9] Sokoloff, D. R., Sanchez, A., Osgood, R. M., and Javan, A., Appl. Phys. Letters **17**, 257 (1970).
[10] Matarrese, L. M., and Evenson, K. M., Appl. Phys. Letters **17**, 8 (1970).

DISCUSSION

K. M. Baird: You tempted us with the picture of the Josephson dewar for seeing how high a frequency response you could get, but you didn't say how high it might be.

K. M. Evenson: We have obtained—or I should say Don McDonald in conjunction with us has obtained—responses up to about 6 to 8 terahertz. We have seen the fifth step from the HCN laser, and we have seen the first step of the 118-μm line of the water vapor laser. There are also some possibilities that the Josephson junction will work excellently as a mixer while it may not actually oscillate at this high a frequency. These experiments are also underway.

A. Javan: I'm sorry. Terahertz. I still have difficulty. What wavelength would that be?

K. M. Evenson: Eight terahertz is roughly 50 microns.

A. Javan: Fifty microns? Fine. Very good.

K. M. Evenson: I haven't converted you yet to terahertz?

A. Javan: I haven't developed a feeling for it, so I just don't know what I would be talking about if I used terahertz. This is why I have been staying away from it. As a matter of fact, it's a very good unit to use. It's quite good. This measurement that Dr. Evenson referred to at 5 microns, which is presumably the highest frequency that one has done the mixing, is one that has been submitted for publication to Applied Physics Letters nearly a couple of weeks ago. Maybe I can mention here very quickly, since the speed of light measurement is being discussed so thoroughly here, that ever since the early days at M.I.T. when we attempted the frequency measurements of lasers, we have had an interferometer to measure the wavelength of a far infrared laser and/or infrared laser accurately. The main purpose of it—really I must confess it is not as much measuring the speed of light as it is measuring the wavelength in the infrared or far infrared—is to improve the wavelength measurement all the way to a part in 10^8. Of course, if one does a wavelength measurement to a part in 10^8, then one has the speed of light. But then again the main emphasis is to have some way of measuring wavelength with that high accuracy, and we have an interferometer to do the job. We have a Michelson interferometer. We compare the fringes of an infrared laser with fringes that are simultaneously observed at 6328 A of helium-neon, and in turn we compare the helium-neon laser with the krypton standard. This has been coming along over the years. In fact, a couple of years ago we had a measure of speed of light which was not any improvement over what had been done but one quite in agreement with it to parts in 10^6. Two parts in 10^6 actually. Maybe I could quickly mention that we have now switched the experiment to 10 microns to compare the 10 micron wavelength with the 6328 A wavelength, and I am pretty certain that we will have a part in 10^8 shortly. As a matter of fact, we are not able to make a comparison to a part in 10^7, but the problem is the drift in the laser within a few megacycles. But then we are in the process of stabilizing lasers, and so on and so forth. And at a part in 10^8 we plan to start on the ten microns. The wavelength comparison by itself, despite the fact that the frequency measurement is so important, is going to have its place. And, in fact, in some ways, you know, if you can do a part in 10^7 measurement of wavelength, directly, with the interferometer, you can do it without the need for two lasers or three lasers. So in defense of wavelength measurement, I am all for it, and maybe one can improve the accuracy to a part in 10^8 or even higher, 10^9, or 10^{10} or better. So both of them have their points.

Determination of the Velocity of Light

A. J. Leikin, V. S. Soloyov, I. V. Lukin, and S. V. Sikora

Kharkov State Institute of Metrology, U.S.S.R.

Measurements of the velocity of light in vacuo were carried out at KhGNIIM in 1958–1966 under the direction of Dr. G. S. Simkin, using the interference of millimeter waves. A two-beam interferometer operating in the distant field region of wave propagation was used for this purpose [1, 2].

The wavelength was measured by a well-known method of optical multiplication (Vaisala's method) over a 4.5 m base line. Refractive indices of air and water vapours were determined by means of a resonant refractometer and a correction was made for temperature gradient along the path of wave propagation.

The calculations of diffraction corrections were carried out by a computer.

During the measurements, which lasted for a long period of time, all precautions were taken against external disturbances of different kinds.

Preliminary experiments (in 1964) gave:

$$c_0 = 299792.5 \pm 0.17 \text{ km/s}.$$

After an improvement of the apparatus and experimental techniques in (1966) we obtained:

$$c_0 = 299792.56 \pm 0.11 \text{ km/s}.$$

Further work on the measurement of the velocity of light is directed towards achieving an accuracy of the order of 10^{-8}. The possibility of achieving such an accuracy is connected with:

—using shorter wavelengths to reduce errors due to diffraction;

—conducting measurements directly in vacuo;

—using laser interferometry for wavelength measurements.

At present preparations are in progress at KhGNIM for light velocity measurements with a laser operating in the infrared. For this purpose it is suggested to use HCN lasers ($\lambda = 337$ μm), H_2O lasers ($\lambda = 118$, 84 and 27 μm) as well as CO_2 lasers ($\lambda = 10.6$ μm).

The main questions to be solved in these experiments are:

—measurements of the frequency of laser emission

—measurements of wavelengths with a correction for diffraction.

At the present time an apparatus is available for the measurement of frequency of a HCN laser ($\lambda = 337$ μm) [3], commercial silicon diodes serving as mixers.

The work on frequency measurements of lasers with $\lambda = 84$ μm, 27 μm and 10 μm is in progress.

Wavelength measurements of lasers are being conducted with an interferometer the maximum path difference of which is 3.5 m by comparing IR-waves with the emission of a He–Ne laser ($\lambda = 0.6$ μm). The red line in its turn is specified against the orange ^{86}Kr-line by means of a 100 mm Fabry-Perot interferometer.

The measurements are carried out by the determination of the fringe number correlation on the path of 3.5 m (Doppler interferometer) as well as by measuring this base by means of a reversible counter.

Two principles are used in the optical interferometer: a usual Michelson interferometer and a laser interferometer with the feedback.

The assessment of the diffraction corrections is made with a computer.

The interferometric part of the experimental setup is placed in a vacuum camera.

The experiments are suggested to be carried out in two stages:

(1) Preliminary measurements to check the apparatus and to obtain an accuracy of 10^{-7} (in air).

(2) Measurements in vacuo to achieve accuracies of the order of $(2 \text{ to } 5)/10^8$.

Parallel investigations are being conducted on the use of He–Ne lasers ($\lambda = 0.633$ μm) with synchronized oscillations for light velocity measurements [4]. The main problems here are those of stabilization of the frequencies of intermode beatings and the wavelength measurements of neighbouring frequencies with an error of $2/10^8$. The expected accuracy is $3/10^8$.

Key words: Doppler and Michelson interferometers; H_2O, HCN, CO_2 lasers; He–Ne lasers; ^{86}Kr; laser interferometry; velocity of light.

References

[1] Simkin, G. S., Lukin, I. V., Sikora, S. V., Strelenskii, V. E., Izmeritelnaya Technika, No. 8 (1967).

[2] Simkin, G. S., Lukin, I. V., Sikora, S. V., Proceedings of KhGNIIM, 2nd issure (1969).

[3] Leikin, A. J., Fertik, N. S., Solovyov, V. S., Bondarev, V. N., Proceedings of the First All-Union Conference on Metrology (Sept. 1968).

[4] Leikin, A. J., Solovyov, V. S., Izmeritelnaya Technika, No. 4 (1970).

RYDBERG CONSTANT

Atomic spectroscopists pine
For a Rydberg with integers nine,
And hope that George Series
Himself never wearies
Of playing his game, "What's My Line?"

ANON.

Determination of the Rydberg Constant

G. W. Series

J. J. Thomson Physical Laboratory, Reading, England

This paper reviews the present situation concerning the Rydberg. The value recommended in recent evaluations of constants rests mainly on measurements of wavelengths of H and He^+ lines made over 30 years ago and is subject to systematic errors arising from incomplete resolution of fine structure components. The very thorough studies of the fine structure which have been made in recent years, and advances in spectroscopic technique, have prepared the ground for a new determination of the Rydberg with substantially improved precision. A concrete recommendation is made.

Key words: Atomic constants; hydrogen fine structure; Rydberg constant.

1. Introduction

In evaluation of the fundamental constants the value of the Rydberg has generally been taken as one of the cornerstones: the uncertainties of measurement are supposed to be so small that the value can be treated as exactly known. Indeed, the value *is* reliable at the level of one part per million, but it is not at all reliable at the level of one part in ten million, which is the uncertainty which has been assigned to it in recent assessments.

It may surprise some of the participants at this Conference to learn that the measurements on which the present value of the Rydberg is based were all made over 30 years ago [1, 2, 3]. The first of these is dated 1927; the second and third, 1939 and 1940 respectively. All these measurements were made at a time when the primary standard of length was the meter bar and the spectroscopic standard the cadmium red line, both of which are now obsolete.

It is not simply on account of its age that we should cast a critical eye upon this work, but we are obliged to ask ourselves whether or not the advances in knowledge which have been made in the intervening period might shed some new light on these old measurements, and whether advances in technique might not offer the opportunity of making a more reliable measurement. I think it is true to say that the old measurements simply cannot survive a critical scrutiny in the light of modern knowledge. They are subject to systematic errors, some of which their authors were aware of, and some of which they were ignorant of. Any assessment which purports to extract from these old measurements a precision of 1 in 10^7 represents an averaging, not simply over random errors, but also over incompletely analyzed systematic errors.

I do not propose here to discuss this old work in detail but I shall try to show what the difficulties are, and how we are now in a much better position to meet them. I am convinced that an improvement in the value of the Rydberg is feasible at the present time, and I should like to state a case in support of this conviction.

2. Definition of the Rydberg Constant

Rydberg introduced his constant into physics in an empirical formula to represent term values for alkali atoms [4]. The expression $N_0/(m+\mu)^2$, with m taking successive integral values, gave a good representation of the experimental term values in a particular series. μ was a nonintegral constant, different for different series, and N_0 was a universal constant whose value was 109721.6 cm^{-1}.

Rydberg immediately recognized that his expression, with $\mu=0$, was equivalent to Balmer's formula for the wavelengths of lines in the spectrum of atomic hydrogen, but more than 20 years elapsed before Balmer's formula was given a basis in the theory of atomic dynamics by Bohr [5]. To him we owe the relation

$$Rhc = 2\pi^2 me^4/h^2, \tag{1}$$

where R is a wavenumber per unit length, m is the mass of the electron, and its charge, $-e$, is given in esu. h is Planck's constant and c the velocity of light. R is to be identified with Rydberg's N_0.

As we all know, Bohr's formula gives an excellent description of the gross structure of the spectrum of atomic hydrogen, but it does not account for its finer details. Nevertheless, the group of atomic constants which emerges in his theory is a very important combination which may, not without difficulty, be related to the wavelengths of spectral lines of one-electron atoms, and in turn to the primary standard of length. It is this combination of constants which we define to be the Rydberg Constant:

$$\begin{aligned} R &= me^4/4\pi\hbar^3 c \qquad \text{(cgse)} \\ &= (\mu_0 c^2/4\pi)^2 me^4/4\pi\hbar^3 c \qquad \text{(SI)}. \end{aligned} \tag{2}$$

By this definition the units of R are $(\text{length})^{-1}$, but this is only a matter of experimental convenience, reflecting the fact that it is easier, in the visible region of the spectrum, to measure the wavelength of a spectral line than to measure its frequency. If it becomes feasible to measure the frequencies of spectral lines in the visible, then it would be appropriate to redefine the Rydberg in units of $(\text{time})^{-1}$.

In passing, it is worthwhile recalling other combinations of constants in terms of which the Rydberg may be expressed:

$$R_E = \tfrac{1}{2}(me^4/\hbar^2) = \tfrac{1}{2}(e^2/a_0) = \tfrac{1}{2}mc^2\alpha^2, \tag{3}$$

where R_E is in units of energy, a_0 is the Bohr radius, and α the fine structure constant.

R occurs naturally in combination with powers of the fine structure constant in theoretical expressions for the fine and hyperfine intervals in hydrogen. Thus, the Dirac formula gives the fine structure intervals in terms of $R\alpha^2$, the Lamb shift constant is proportional to $R\alpha^3$, while the hyperfine intervals are again proportional to $R\alpha^2$. Because of uncertainties in α and in other constants, these intervals are not so suitable for the determination of R as is the gross structure.

3. The Determination of R by Measurement of Wavelengths

Up to the present time the determination of R has been based on the measurement of wavelengths of selected lines in the spectra of H, D, and He^+. The existence of fine structure, which we shall discuss at length, introduces a serious difficulty into this work. An important, but much more straightforward point arises in connection with the correction to Bohr's formula for the finite mass of the nucleus. The corrected "Bohr" wavenumbers for an atom of nuclear mass M and atomic number Z are given by

$$\tilde{\nu}(n_1, n_2) = R_M Z^2 (1/n_1^2 - 1/n_2^2) \tag{4a}$$

where

$$R_M = R_\infty (1+m/M)^{-1}. \tag{4b}$$

The R_∞ of this equation is the R which we defined in eq (2). R_M is the quantity obtained from experiment when the corrections for fine structure have been made. The factor $(1+m/M)^{-1}$ is needed to allow comparisons between the Rydberg obtained from one-electron atoms of different mass. This factor is known to a precision of about 1 in 10^8 [7], which is adequate at the present level of precision, but which would introduce an uncertainty which could not be entirely ignored if the improvements suggested in this review were realized.

The fine structure is represented by terms in the theory additional to eq (4). The most important of these terms are given by the Dirac theory which yields, for the energy of the state (n, j) of a one-electron ion of atomic number Z, the expression

$$\frac{E_{n,j}}{hc} = \frac{-R_M Z^2}{n^2}\left[1+\frac{\alpha^2 Z^2}{n}\left(\frac{1}{j+\frac{1}{2}}-\frac{3}{4n}\right)\right.$$
$$\left.+(\text{terms of higher order in } \alpha^2 Z^2)\right]. \quad (5)$$

(The expression can be written in closed form so that no inaccuracy need be introduced here.) There is an additional small term which depends only on n, and "Lamb shift" terms which remove the degeneracy in l. The Lamb shift is greatest for s-states. Figure (1b) shows the fine structure for the levels $n=2$ and 3 in hydrogen. Values of the energy, including all these terms, have been very conveniently tabulated by Garcia and Mack [6].

The determination of the Rydberg, then, rests on measuring the position of suitable fine structure components and applying a correction to bring the measured value into the "Bohr" position given by eq (4). Such corrections are generally computed from the theory. It is not possible, of course, to make a direct experimental test of the *displacements* of the levels from the "Bohr" positions, but many of the predicted *intervals*, that is, differences between displacements, have been verified experimentally to a precision which is ample for the present purpose. The corrections computed from the theory may therefore be used with confidence. The largest displacement is that of the level $1^2S_{1/2}$, and this we shall discuss separately.

The difficulty concerning the fine structure arises because this structure has not been resolved in any of the determinations upon which the value of the Rydberg is based. There are a variety of reasons for this lack of resolution, but the dominant theme is that the Doppler width of individual components obscures the structure. If the structure could be resolved to the extent that the precision of measurement was limited only by the techniques of interferometric spectroscopy, then it would be possible to improve the value of the Rydberg by nearly an order of magnitude. The corrections necessary to bring the measured position of any fine structure component into the "Bohr" position may be taken as very reliably known.

In hydrogen there is the further complication of hyperfine structure, and in some future determinations it may be necessary to take this into account. At the present time the hyperfine structure has not been resolved in the optical spectrum. The magnitude of the splittings are reliably known from microwave spectroscopy. An account of the hyperfine structure may conveniently be found in the book or encyclopedia article by Bethe and Salpeter [8].

4. The Experimental Basis of the Present Value

As an illustration of the experimental problem I should like to refer to the work of Drinkwater, Richardson and Williams [3] in one of the investigations on which our present value of the Rydberg is based. (I take as the present value the number given in the recent review by Taylor, Parker, and Langenberg [7].) Drinkwater and his collaborators measured the wavelengths of certain components of the first Balmer line (hydrogen alpha; 6563 Å; $n=2$–3) in hydrogen and deuterium. The light source was a gas discharge cooled by liquid air. The spectra were recorded photographically, and the cadmium red line was used as a standard of wavelength.

Figure 1a shows a photometer tracing of one of their spectrograms of H_α. These are fringes obtained with a reflection échelon, and need correcting for the strongly wavelength-dependent instrumental function.

Figure 1b is a term diagram based on the theory as we now have it, i.e., including the Lamb shifts. It is clear that if figure 1a is to be interpreted on the basis of 1b, some drastic assumptions need to be made concerning the relative intensities of the

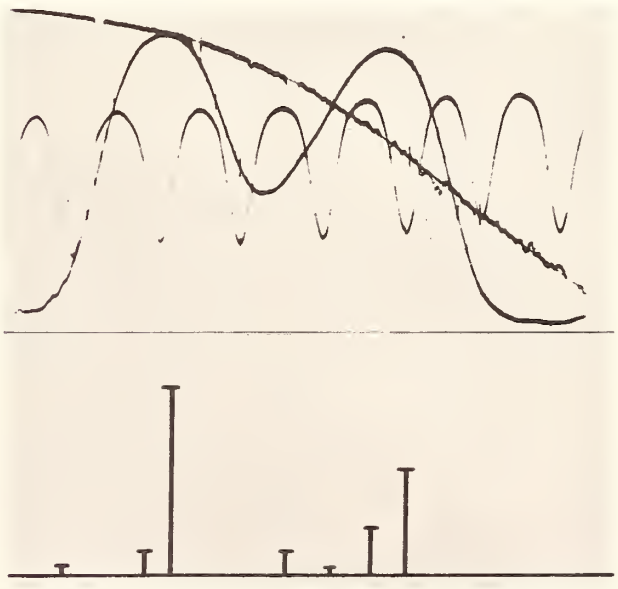


FIGURE 1a. *Fine structure of* H_α *from a discharge cooled with liquid air* [3].

The M shaped trace is H_α. The trace showing oscillations is for calibrating the wavenumber scale; the trace falling from left to right is for calibrating the intensity.

Below the curves is the theoretical fine structure arising from the transitions shown in fig. (1b).

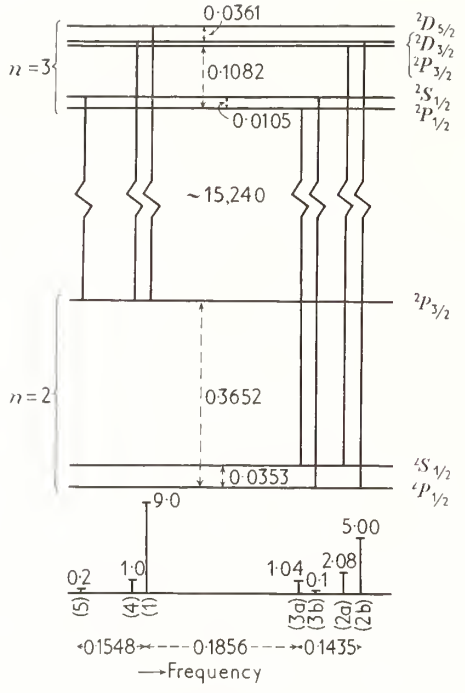


FIGURE 1b. *Theoretical fine structure, including Lamb shifts.*
Intervals are in cm^{-1}. Relative transition probabilities are marked on the components.

components, the absence of false lines, and whether the conditions in a gas discharge correspond to those of the free atoms to which figure 1b relates. The uncertainty of 1 part in 10^7 which is claimed for the present value of the Rydberg implies an uncertainty of 0.0015 cm^{-1} in the wavenumber of H_α, which is about 1/200 of the interval between the main peaks in figure 1a. It is surely over-optimistic to expect precision of this order in the analysis of 1a into a structure as complicated as 1b, and as open to perturbing influences.

As an indication of the unreliability of analyses of such incompletely resolved structures we may recall that there was some controversy about the position of the component $2^2S_{1/2}$–$3^2P_{1/2}$ at the time when Drinkwater, Richardson, and Williams were reporting their measurements. These authors concluded that they found no discrepancy with the predictions of the Dirac theory. But the Lamb shift displaces this component from its position as given by the Dirac theory by 0.03 cm^{-1}, and further, the components $2^2(S, P)_{1/2}$–$3^2(P, D)_{3/2}$, which are degenerate in the Dirac theory, are split by the Lamb shift into two components separated by 0.03 cm^{-1}. This figure is 20 times as great as the uncertainty which is claimed for the absolute wavenumbers of the components of H_α when 1 part in 10^7 is claimed for the Rydberg. Yet neither the shift nor the splitting were detected by Drinkwater, Richardson, and Williams.

This study of H_α has been selected to illustrate the problem of analyzing the spectra in relation to the precision which is claimed. Similar problems were met by the other investigators in the analysis of H_β and worse problems in the case of the He^+ line 4686 Å. In no case was the structure resolved and no case is free from the possibility of systematic error at the level of parts in 10^7.

It should be added that the old measurements have been reassessed in the light of our present knowledge of the Lamb shift [9], and that the present recommended value of the Rydberg has been arrived at in the light of this reassessment. But more than a reassessment is needed. No such reassessment can make proper allowance for intensity anomalies, or for such technical considerations as the non-uniform distribution of light in the image plane of the optical system and the crudeness of photographic photometry as compared with photoelectric methods.

5. Choice of Line: the Fine Structure

In considering which spectral lines are most suitable for a determination of R there would appear to be no advantage gained in looking beyond the spectra of the isotopes of H and He^+. Within these spectra the simplest lines are the resonance lines, $n=1$–2. These offer considerable attractions but also present special difficulties since they lie in the vacuum ultraviolet. We shall defer a discussion of these and consider first those lines which lie in the visible and near ultraviolet.

The simplest line of hydrogen which lies in the visible is Balmer-α, 6563 Å, $n=2$–3. The simplest line of He^+ is the Paschen line, 4686 Å, $n=3$–4. These are the lines which have normally been studied, but higher series members have also been investigated.

The structures of the H and He^+ lines become more complicated as one goes to higher series members. Their overall width is substantially determined by the sum of the $^2P_{1/2,3/2}$ intervals of the upper and lower states. Within this spectral region are numerous components (7 for Balmer-α, 13 for the Paschen line) with a great diversity of intensities. In the Balmer series the components tend to cluster in two groups, separated by the $2^2P_{1/2,3/2}$ interval, so that with incomplete resolution most of the Balmer lines appear as doublets.

In this situation one might attempt to resolve as completely as possible one or more components and concentrate the effort of measurement upon these, or one might give up any attempt at resolution and try to measure the centers of gravity of blends. I believe that the former method is strongly to be preferred and that the latter method is open to systematic errors for reasons to be set out below. Nevertheless, it must be admitted that this latter method has given remarkably consistent results. In a recent series of measurements, Csillag [10] has obtained 12 values of the Rydberg from the six doublets D_β–D_η. The lines were excited by a high-frequency discharge in a liquid nitrogen cooled discharge at low pressure. The values are self-consistent to within 1 part in 10^7 and agree with the best value obtained from the earlier measurements.

It appears to me that the value of this work lies in the evidence it provides that hydrogen atoms may be excited in a gas discharge remarkably free of perturbation. The higher states of hydrogen are more susceptible to the Stark effect than the lower, and if no evidence of displacements is found for the higher series members, then one can feel confident that, to this level of precision, neither the higher levels nor the lower are appreciably perturbed. Indeed, the stimulus to Csillag's investigation was to determine whether the theory of fine structure was adequate for the higher terms in hydrogen.

A more recent determination of the Rydberg from measurements of the H_α line will be reported at this conference [11]. The use of the light isotope in this work, with liquid nitrogen cooling of the discharge, implies considerable overlapping of the components. The interferometric technique is different from that which earlier investigators have used. The position of the components is found to depend on the discharge current. Extrapolation to zero current yields the value $R_H = 109\,677.593 \pm 0.0035$ cm^{-1}, which leads to a value of R_∞, rather higher than the value accepted at present.

5.1. Anomalies in the Fine Structure

It is appropriate at this point to give an account of anomalies which have been found in studies of the fine structure, and to show how these anomalies can lead to systematic errors in measurements of unresolved blends.

Nearly all the older studies of the fine structure in H and He^+ reported that the relative intensities of the components did not agree with the relative transition probabilities calculated from the Dirac theory. This is not evidence against the theoretical transition probabilities, but reflects *either* a nonstatistical distribution of atoms over the excited states, *or* that the transition probabilities are being perturbed by the fields in the discharge tube. The intensity ratios have often been found to depend on the discharge current. It is clear that there will be systematic errors in the value of the Rydberg obtained from unresolved structures if the centers of gravity of blends are calculated from theoretical intensity ratios which are known to be unreliable. Moreover, if the anomalous intensities are attributable to electric fields in the discharge, one would expect to find Stark shifts in the positions of the components. It is important, therefore, to try to discover the origin of any intensity anomalies which might be observed.

The possibility of shifts and intensity anomalies arising from electric fields in the discharge has been considered by most authors who have studied the fine structure. The theoretical analysis of the Stark effect in hydrogen is, of course, very well known [8]. Certain components are more sensitive than others to electric fields, and one can set upper bounds on the fields present in any particular circumstances by studying the relative intensities or displacements of these components. Such studies are the more valuable, the more completely the components are resolved. In this context, the recent studies of the He^+ 4686 line by Roesler and Mack [12] and by Berry and Roesler [13] are most instructive.

By studying the relative intensities of components of spin doublets (e.g., $3^2S_{1/2}$–$4^2P_{3/2,1/2}$) these authors were able to set an upper bound as small as 10 V/cm on the electric field in their hollow cathode plasma. Evidence from this and other investigations shows that the Stark effect on atoms in gas discharges can be kept negligibly small.

The components which one would select for a determination of the Rydberg are not those which are most sensitive to the Stark effect. A study of the components which *are* sensitive would enable one to make a realistic assessment of the systematic error arising from electric fields, and to reduce this error to acceptably small limits.

Intensity anomalies associated with displacements of components of the fine structure can arise for other reasons. Again, the He^+ line 4686 Å furnishes an example. Some years ago [14] I obtained good resolution of the structure of this line by using a discharge in a hollow cathode cooled by liquid hydrogen, but the components which arose from the $S_{1/2}$ level of the upper state were anomalously strong. I also found a small disagreement with theory in the position of one of the components. It was subsequently shown by Herzberg [15] that the relative intensities depended on the pressure of the gas, and that the intensity anomalies were attributable to nonstatistical populations of the excited states. For the higher pressures, statistical populations were achieved, but at the lower pressures the populations depended on the different excitation probabilities and lifetimes of the different states. A theoretical analysis has been given by Bethe and Salpeter [8].

A very subtle source of systematic error is now revealed. Granted that ions in different fine structure levels of the excited state have different mean lifetimes, it follows that they will have different mean drift velocities in a dc discharge, and this will lead to a differential Doppler shift of the fine structure components. This effect was systematically studied by Roesler and DeNoyer [16] and explains the discrepancy which I had found in the line position. Here again an intensity anomaly is intimately connected with the systematic displacement of a component from its theoretical position. This effect would not arise in the fine structure of neutral atoms.

5.2. The Vacuum Ultraviolet

The effects here discussed for lines excited in gas discharges would not arise for lines studied in absorption, and for this reason the Lyman series has a great attraction. Moreover, the first member, Lyman-α, has a very simple fine structure: a doublet, with intensity ratio 2:1. Unfortunately, both for helium and for hydrogen, the line lies in the vacuum ultraviolet, where the techniques of interferometric spectroscopy have not yet been developed although Freeman at the N.P.L. is making substantial progress in this direction. The wavelength of this

line in deuterium (1215 A) was in fact measured by Herzberg [17] with a view to measuring the Lamb shift of the ground state. He used a vacuum grating spectrograph and was able to determine the wavelength to about 3 parts in 10^7. He suggested that improvements in the technique seemed possible which would bring this figure down to about the same precision as the value of the Rydberg. While this improvement would be valuable in checking the Lamb shift of the ground state it does not seem likely that techniques based on grating spectroscopy will be adequate for improving the value of the Rydberg.

It would be wrong to dismiss Lyman-α completely in this connection, since it is not out of the question that interferometers will be developed for the vacuum ultraviolet. There are indeed problems in applying interferometric techniques to absorption spectroscopy, but these have been overcome in the visible by the method of circular channels suggested by Treanor [18] in the case where the background is formed by a continuous spectrum. Alternatively, the background may be formed by a broadened emission line (in this case, from a Lyman-α source). The Doppler width of the absorption line could be reduced sufficiently to allow resolution of the doublet by using a collimated atomic beam. However, a point of principle remains: the fine structure "correction" for Lyman-α includes the Lamb shift of the ground term $1^2S_{1/2}$ which is the largest Lamb shift in the spectrum, and not open to independent experimental study. Any determination of the Rydberg from the wavelength of this line would be critically dependent on the theoretical value of the Lamb shift.

This objection does not apply to the He^+ line 1640 Å ($n=2$–3), which corresponds to Balmer-α. This line was studied by Herzberg [15] who resolved the fine structure more completely than has been done for Balmer-α in hydrogen: (the scale of the Dirac terms in the fine structure is sixteen times as great for He^+ as for H). This again was achieved with a grating spectrograph. It is probable that still more complete resolution could be obtained by exciting the line in a liquid helium cooled discharge, and by correspondingly improving the spectroscopic resolving power. Against the attractions of a resolved structure must be set the difficulties of making precision measurements of wavelength in the vacuum ultraviolet and the difficulties already mentioned associated with the excitation of He^+ lines in a gas discharge. Moreover, as is pointed out below, there is no advantage to be gained in choosing an ultraviolet line rather than one in the visible so long as the linewidth is determined mainly by the Doppler effect.

6. Spectral Line Width

I have argued for the desirability of resolving the fine structure components: I must now discuss the practicability.

The dominant cause of line broadening in all investigations of the fine structure of hydrogen and ionized helium has been the Doppler effect, and all advances in resolving the fine structure have been based on reducing this width.

It should first be pointed out that the Doppler width of a spectral line is proportional to its frequency, so that, insofar as the precision of measurement is limited by this factor, there is no advantage in choosing a high-frequency spectral line rather than one of lower frequency.

6.1. Excitation in a Gas Discharge

For lines excited in a gas discharge, where the radiating atoms are at temperature T K, the Doppler width at half intensity, $\Delta\nu_{1/2}$, is given by

$$\Delta\nu_{1/2}/\nu = 7.2\times10^{-7}(T/M)^{1/2} \qquad (6)$$

where M is the molecular weight. One immediately sees the reason for cooling the discharge, and for choosing deuterium rather than hydrogen. The temperature of the emitting atoms is not, of course, the same as the temperature of the cooling bath; nevertheless, by careful design of the discharge tube and by running the discharge as gently as possible one can achieve very substantial reductions of temperature. The following figures, obtained from analyses of spectral profiles, have been reported:

(a) 40 K: D_α line from discharge in deuterium cooled by liquid hydrogen under reduced pressure [19]

(b) 35 K: He (neutral atom) lines from discharge in He cooled by liquid helium [12]

(c) 100 K: He^+ 4686 Å line in the same discharge

(d) 4–11 K: He (neutral atom) lines from discharge in He cooled by liquid helium under reduced pressure [20].

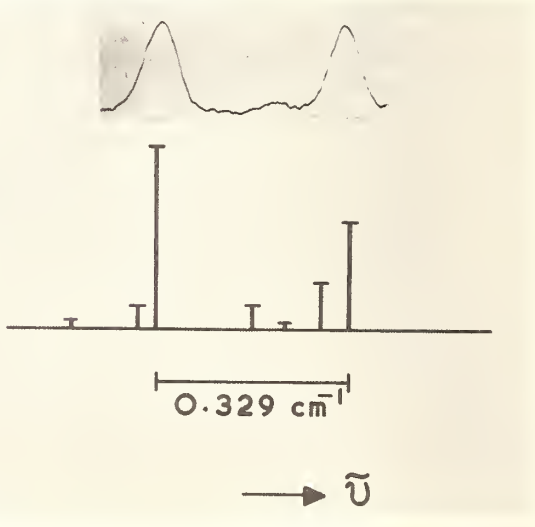


FIGURE 2. *Fine structure of D_α from a discharge cooled with liquid hydrogen at reduced pressure [19].*
The improvement in resolution over the curve in fig. (1a) is due partly to the use of the heavier isotope, partly to the colder bath, and partly to the use of a milder discharge.

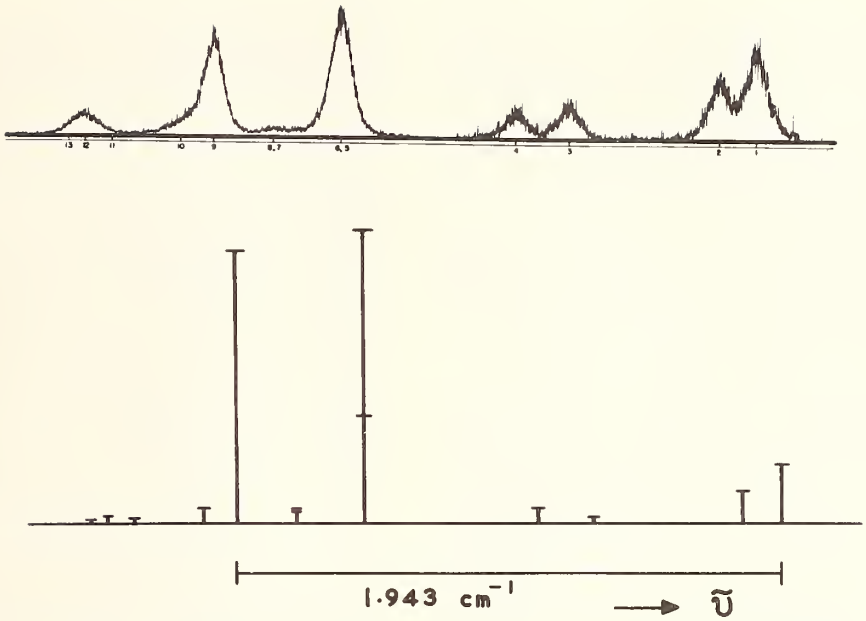


FIGURE 3. *Fine structure of* He^+ *4686 Å from a discharge cooled with liquid helium* [12].
(The gain has been increased by a factor 3 for the four high frequency components). The theoretical fine structure pattern shown below the curve includes the Lamb shifts. The excellent resolution of practically all the components is to be noted. In this recording there are intensity anomalies.

In figures 2 and 3 we see recordings of spectra from which some of these figures were obtained. It is clear that the major components are well resolved, and would require only slight corrections for minor, overlapping components. In both cases these studies were undertaken to investigate the fine structure, not to measure the Rydberg.

The width in example (c) appears anomalous by comparison with (b) and requires discussion. A very detailed study of the width of individual components of the He^+ line was made by Berry and Roesler [13], who interpreted the broadening in terms of the detailed mechanism whereby He^+ states are excited by electron impact in the discharge. The recoil of the ions introduces an excess Doppler broadening of the line. This interpretation was supported by a comparison of the broadening in $^3He^+$ and $^4He^+$. This broadening appears to be an ineseapable feature of excitation by electron impact.

6.2. Excitation in Atomic Beams

a. The Paschen Line

A different approach to the problem of producing the 4686 Å line with narrow components was undertaken by Larson and Stanley [21]. These authors excited the line by electron bombardment of a highly collimated atomic beam of He. This technique had been proved to be very successful for many lines of the rare gases [22]. In this method of excitation the components of the 4686 line appeared with very different relative intensities from those found in a gas discharge. Nevertheless, the line widths were very similar, and were explained in this case also in terms of the dynamics of the collisions.

Some details of the limitations imposed on the line width by the mechanism of excitation by electron impact will be discussed in a contributed paper [23], and other possibilities for exciting the line will be proposed.

b. The Balmer Lines

An attempt was recently made by Stoner [24] to reduce the linewidth of the components of the Balmer lines by exciting them by electron bombardment of a highly collimated beam of atomic hydrogen. The light was analyzed by a Fabry-Perot étalon and detected by pulse-counting techniques. While a substantial improvement in linewidth was achieved, the signal-to-noise ratio was extremely unfavorable on account of the low intensity of the emitted light. No useful precision measurements could be made.

6.3. A New Proposal for the Balmer Lines

I should like to make a suggestion for reducing the linewidth of the Balmer lines—in particular the first member—by a different technique, namely, by attempting to excite the line in a gas discharge cooled by liquid helium. The hydrogen would, of course, solidify at liquid helium temperatures, but helium could be used as a carrier gas in the discharge in much the same way as rare gases are used as carriers in alkali metal lamps. Molecular hydrogen has an appreciable vapor pressure in the range 7–10 K so that one might reasonably expect to see the atomic spectrum with rising temperature in the helium discharge. Cooling would be most effective in the positive column, but if there were any difficulty in exciting the spectrum one might try a hollow cathode.

Moreover, the use of tritium would secure a further narrowing of the lines. The fine structure of tritium-α has already been studied [25], but the excitation was in a discharge tube cooled by liquid nitrogen.

7. Technical Considerations Relevant to a Re-Determination

Having discussed the nature of the lines whose wavelength we might wish to measure, we now turn to the primary standard of length and then consider briefly some points of spectroscopic technique.

7.1. The Primary Standard of Length

The meter has, since 1960, been based on the vacuum wavenumber, 1 650 763.73 m^{-1}, for the radiation connecting the unperturbed levels $2p_{10}$ and $5d_5$ of ^{86}Kr [26]. Krypton lamps, prepared to a recommended specification, are maintained in the various standards laboratories. The orange line which corresponds to the primary standard has been studied very carefully under variations of pressure, current density, and temperature, and it is believed that the standard can be actually realized and reproduced to a precision not worse than 1 part in 10^8. This is a conservative estimate.

There is a small difference in wavelength depending on whether one takes the light travelling in the direction from cathode to anode or vice versa: preference is given to the former. The line profile is dominantly Gaussian, with a width about 13 mK. There is some evidence that it is slightly asymmetrical.

The sharpness of the krypton line, given by the ratio of $\nu/\Delta\nu_{1/2}$, is about 1.3×10^6. The corresponding figure for the hydrogen and ionized helium lines illustrated in figures (2) and (3) is, in each case, about 2.5×10^5. The "Q" of these lines therefore approaches that of the standard of length, as maintained, to within a factor 5.

7.2. Spectroscopic Technique

Comparison of wavelengths at the highest level of precision is invariably by some interferometric method. In the visible region of the spectrum the Fabry-Perot étalon is most frequently used, though Fourier transform spectroscopy based on use of the Michelson interferometer is also practiced. These instruments have substantial advantages in light gathering power.

Much experience in their use has been gained in recent years, particularly in the use of double Fabry-Perot étalons to obtain good resolving power combined with extended spectral range. This is particularly important in connection with the He^+ line 4686 Å, since the width and complexity of the structure preclude resolution of the components with a single étalon. In Chu's investigation [2], for example, where a single étalon was used, components from different orders overlapped very considerably, whereas the spectrum shown in figure 3 where the components are resolved was obtained with a double étalon.

Reflecting layers formed of multiple dielectric films are much used in modern interferometry, but they are unsuitable for intercomparisons of wavelength because of the strong wavelength-dependence of amplitude and phase change on reflection. Metal films are generally to be preferred. There is a dispersion of phase change with metal films, but this can be determined and allowed for by the use of spacers of different lengths.

When wavelengths are compared in air it is necessary to take account of the dispersion of refractive index. This is known with an accuracy more than adequate for making the correction [27]. Since the primary standard is defined as a wavenumber in vacuo the wavenumber in vacuo of the eomparison line may be determined without requiring knowledge of the refractive index itself.

The big advance in interferometry since the earlier determinations of the Rydberg has been in the use of photoelectric recording. It was formerly the practice to photograph fringes and to measure their positions visually, using a travelling microscope. Line profiles were studied by making microdensitometer traces and calibrating the optical density by means of intensity standards. Profile analysis was difficult and unreliable.

Photoelectric recording is generally performed by measuring the flux of light which enters a small aperture at the center of the fringe system. The fringes are scanned by changing the optical path in the étalon, either by varying the pressure of gas in an enclosed chamber or by controlled movement of the plates. The signal-to-noise characteristics of this method are good, and since the photoelectric signal is accurately proportional to the light intensity, the fringe profile is reliable and can be subjected to exhaustive analysis.

It is to be noticed that the method of center-spot scanning requires only a small illuminated area, in contrast to the photographic method which requires that the illuminated area should embrace several fringes. The projected area of the light source can therefore be made smaller if the optical components are otherwise similar and the source itself can be designed to dissipate less energy and to be more effectively cooled.

7.3. Other Considerations

The precision of wavelength measurement depends on the signal-to-noise ratio no less than on the sharpness of the lines, and in this respect the hydrogen-α line has a distinct advantage over the He^+ line 4686 Å. The latter is a weak line in the helium discharge and needs relatively high current densities for its excitation. H_α, on the contrary, is the strongest line in the visible spectrum of atomic hydrogen and can be excited in very mild discharges. The gain in luminosity is more than sufficient to compensate for the difference in photomultiplier sensitivities at the two wavelengths. Moreover, the risk of systematic errors due to Stark effect displacements is much reduced with a mild discharge.

There is, further, more risk of contamination by

foreign lines for the He^+ line than for H_α. For the former, one must guard against scattered light from the very strong line of the He arc spectrum at 4713 A. Also, it is known that band lines of molecular hydrogen and of the nitrogen molecular ion overlap the fine structure of 4686 Å, and that a band line due to He_2 falls on the wing. In addition, for the He^+ line, the mechanism which gives rise to the broadening could also give rise to a displacement.

For H_α there is the risk of contamination by the fine structure components of the other isotopes: this is an important consideration when working with deuterium or tritium because of the possibility of exchange with light hydrogen in tap grease in the system. The isotopic displacement arising from the center of mass effect is many times larger than the fine structure splitting, and the fine structure patterns of different isotopes, separated by several orders of interference, overlap in an arbitrary way depending on the length of the étalon spacer. But by suitable choice of spacer the components belonging to different isotopes may be superimposed—the isotope shift is known with sufficient accuracy—and the presence of small amounts of contaminating isotopes rendered innocuous.

For the He^+ line the strong component $[3^2(P, D)_{3/2}-4^2(D, F)_{5/2}]$ would form a natural choice for a determination of R and the weaker components on the shortwave side would serve to monitor the Stark effect. For H_α (we include here D_α and T_α), the strong component $2^2P_{3/2}-3^2D_{5/2}$ would be used for R. It is unlikely that the neighboring weak component $2^2P_{3/2}-3^2D_{3/2}$ could be resolved, but since the two lines belong to the same $D_{3/2,5/2}$ doublet one could be confident about their relative intensities. The other components have not yet been so well resolved as in the case of the He^+ line, but with the improved resolution which I believe to be feasible, these components would similarly be available for monitoring the Stark effect.

8. Conclusion

My thesis is that the fine structures of hydrogen and ionized helium have now been so well explored, and that spectroscopic techniques have been so improved, that it is timely for a redetermination of the Rydberg based upon measuring the wavelengths of spectroscopically resolved fine structure components. At the present time it appears that the long-wave component of tritium-α would be the best line to study. The sharpness of the corresponding component in deuterium-α as measured in fine structure investigations many years ago was already within a factor 5 of the sharpness of the primary standard as now maintained, and there is reason to hope that substantial improvement may be attained with a liquid-helium-cooled light source.

Nevertheless, the primary standard "as maintained" reproduces the primary standard "as defined" only to about 1 part in 10^8. Intercomparisons between lines from the same krypton lamp (which can be used as secondary standards) are reliable to about 2 parts in 10^8. It would therefore seem to me to be over-optimistic to hope to obtain for the Rydberg constant a precision much better than, say, 5 parts in 10^8. The value quoted in the review by Taylor, Parker, and Langenberg [7] is $1.097\,373\,12\,(11)\times10^7$ m^{-1}, that is, a precision of about 1 part in 10^7. It might be argued, therefore, that an improvement in precision by a factor no more than 2 would hardly be worth the effort. I believe, however, that a simple comparison of these figures of estimated uncertainty does not properly represent the situation. Spectroscopic techniques and our understanding of the fine structure have been so much improved since the earlier determinations that a new measurement, skillfully carried out, cannot fail to yield a much more reliable value. I believe that the precision of the currently accepted value may have been overestimated.

It will be appreciated that I have confined my remarks and suggestions to the realm of what is immediately practicable. In a more speculative vein one might contemplate the possibility of using hydrogen atoms in the metastable $2^2S_{1/2}$ state as absorbers in a tunable laser locked to some component of the Balmer-α line. Such a device would allow one to take advantage of the precision in wavelength determination which is offered by the use of lasers of this type.

9. Acknowledgments

Figures 1a and 2, 1b and 3, are reproduced by kind permission of the authors and of the Royal Society, the Oxford University Press, and the American Institute of Physics, respectively. I am grateful to Dr. W. R. C. Rowley of the National Physical Laboratory for a discussion on the primary standard of length and to Drs. F. Ansbacher, H. G. Kuhn, and D. N. Stacey for their comments on the manuscript.

10. References

[1] Houston, W. V., Phys. Rev. **30**, 608–13 (1927).
[2] Chu, D.-Y., Phys. Rev. **55**, 175–80 (1939).
[3] Drinkwater, J. W., Richardson, O., and Williams, W. E., Proc. Roy. Soc. **A174**, 164–88 (1940).
[4] Rydberg, J. R., Vetensk, K. Svenska, Akad. Handl. **23**, (1889).
[5] Bohr, N., Phil. Mag. **26**, 1 (1913).
[6] Garcia J. D., and Mack, J. E., J. Opt. Soc. Am. **55**, 654–85 (1965).
[7] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375–496 (1969).
[8] Bethe, H. A., and Salpeter, E. E., Quantum Mechanics of One- and Two-Electron Atoms. Handbuch der Physik XXXV, 88 (1957). Also separately published, Springer, Berlin (1957); Academic Press, New York (1957).
[9] Martin, W. C., Phys. Rev. **116**, 654–6 (1959).
[10] Csillag, L., Acta Phys. Acad. Sci. Hungar. **24**, 1–18 (1968).
[11] Masui, T., Proceedings of this Conference.
[12] Roesler, F. L., and Mack, J. E., Phys. Rev. **135**, A58–71 (1964).
[13] Berry, H. G., and Roesler, F. L., Phys. Rev. **A1**, 1504–17 (1970).
[14] Series, G. W., Proc. Roy. Soc. **A226**, 377–92 (1954).

[15] Herzberg, G., Z. Phys. **146**, 269–80 (1956).
[16] Roesler, F. L., and DeNoyer, L., Phys. Rev. Lett. **12**, 396–8 (1964).
[17] Herzberg, G., Proc. Roy. Soc. **A234**, 516–28 (1956).
[18] Treanor, P. J., Mon. Not. Roy. Astron. Soc. **109**, 389 (1949). also Adam, M. G., Mon. Not. Roy. Astron. Soc. **112**, 546–69 (1952); **115**, 422–6 (1955).
[19] Kuhn H., and Series, G. W., Proc. Roy. Soc. **A202**, 127–42 (1950).
[20] Brochard, J., Chabbal, R., Chantrel, H., and Jacquinot, P., J. Phys. Rad. **18**, 596 (1957).
[21] Larson, H. P., and Stanley, R. W., J. Opt. Soc. Am. **57**, 1439–49 (1967).
[22] Stanley, R. W., J. Opt. Soc. Am. **56**, 350–6 (1966); also Odintsov, V. I., Opt. i Spekt. **10**, 403. (Opt. Spect. **10**, 202) (1961).
[23] Kessler, Jr., E., and Roesler, F. L., Proceedings of this Conference.
[24] Stoner, J. O., Thesis, Princeton University (1963).
[25] Kireyev, P. S., Dokl. Akad. Nauk SSSR **112**, 41 (1957).
[26] Trans. Int. Ast. U. XIA, 97–117 (1962).
[27] Edlén, B., J. Opt. Soc. Am. **43**, 339 (1953).

DISCUSSION

A. JAVAN: The measurements in gas discharge—What are the pressure ranges when these measurements are done? Could you make some comment on difficulties and collision shifts?

G. W. SERIES: Yes. All these things like collision shifts are small compared with the other causes of broadening. The pressures of hydrogen used in the gas discharge are generally substantially below 1 torr. But one would mix in some helium so that foreign gas would be present as well. With the He^+ line one might go to rather lower pressures, but you may say 1 Torr would be the upper limit for the pressures. That sort of thing.

A New Determination of the Rydberg Constant

Toshiro Masui

National Research Laboratory of Metrology, Tokyo, Japan

The Rydberg constant has been determined from observation of the two-beam interferogram of Hα. The hydrogen source is cooled with liquid nitrogen and excited by 50 Hz ac. Observations were made of the order of interference of Hα in terms of the krypton 86 standard line at path differences up to 3.24 cm and at a discharge current of 10 mA, and of the variation in order of interference at a path difference of 3.24 cm and at discharge currents between 5 and 15 mA.

The final result is $R_H = 109\,677.593\,7 \pm 0.003\,5$ cm^{-1}.

Key words: Balmer series; fine structure; hydrogen; Rydberg constant.

1. Apparatus

The formal procedure for the determination of the Rydberg constant consists of the measurement of a wavelength of an emission from a two-body system composed of a nucleus and an electron. In our new determination, we have made use of a hydrogen discharge tube as the source, a Möbius-band interferometer [1] for the optical system, a photomultiplier as the detector, and a krypton 86 lamp as the reference standard.

A number of hydrogen discharge tubes, appropriate for use in side-on observation of the capillary when immersed in liquid nitrogen, have been constructed, and a single most suitable tube was selected for final use. When the tube is excited by a 50 Hz alternating current of 10 mA at the boiling point of nitrogen, it continues to emit a sufficiently pure and intense atomic spectrum from the central part of the capillary for several hours. At lower currents, the emission of the atomic spectrum breaks down within several minutes and is replaced by a molecular spectrum.

Since the visibility of the two-beam interferograms of Hβ and subsequent lines of the Balmer series at path differences over 1.5 cm has been extremely low, we have finally adopted Hα alone for observation.

The part of the interferometer where the double-mirror reflector moves along the optical axis can be evacuated down to the order of 10^{-2} mm Hg, so that the order of interference in vacuum can be determined directly without correction due to dispersion of air.

2. Order of Interference

At a discharge current of $j = 10$ mA, the order of interference of Hα has been determined in terms of the primary standard krypton 86 line twice at each of the nine path differences indicated in the first column of table 1. Three or four other krypton 86 lines were also included in the observation to permit application of the method of excess fractions. To eliminate the effect of asymmetry between the two arms of the interferometer, we always made use of the sum of the results at two positions of the reflector corresponding to path differences of $\pm D$.

Since the observed interferogram of Hα is the superposition of the two-beam interferograms of all the fine-structure components scattered around the fictitious "Balmer line" [2] whose wave number ν_0 is defined by $R_H(1/2^2 - 1/3^2)$, the observational data without correction represent the quantity $2\nu_0 D + \pi^{-1} \arg F'$, where F' is the Fourier transform with reference wave number ν_0 of the fine-structure pattern of Hα emitted from the actual discharge tube under operation.

Table 1. *Observed order of interference.*

Path difference	Order of interference	Residual error
cm		$\times 10^{-4}$
0.36	10 967.762 4	−64
	765 1	−37
0.72	21 935.548 3	+80
	541 4	+11
1.08	32 903.307 9	−134
	312 4	−89
1.44	43 871.137 8	−39
	147 6	+59
1.80	54 838.703 9	−24
	708 9	+26
2.16	65 806.514 7	−75
	509 4	−128
2.52	76 774.302 1	−1
	314 2	+120
2.88	87 742.069 4	−47
	082 7	+86
3.24	98 709.844 5	+5
	850 2	+62

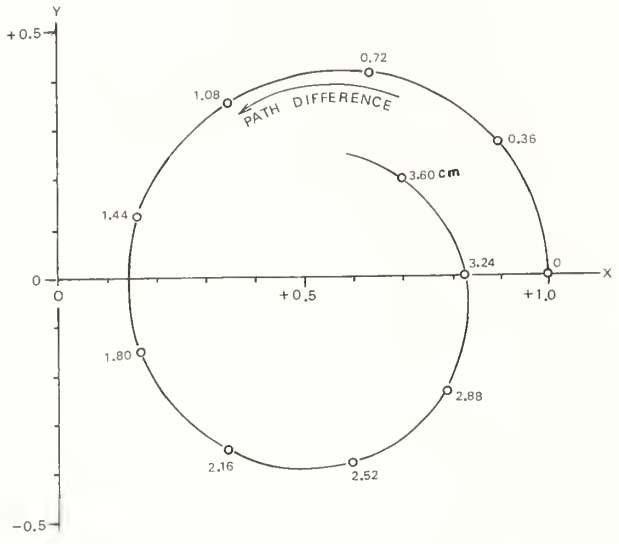


FIGURE 1. *Fourier transform with reference wave number ν_1 of the theoretical fine-structure pattern of* Hα.

The theoretical form of F',

$$F=\sum_k f(\nu_k)\,\exp\{2\pi i(\nu_k-\nu_0)D\},$$

has been evaluated as a function of D, where $f(\nu_k)$ and $\nu_k-\nu_0$ are the theoretically predicted intensity distribution and the splitting of the pattern. Their numerical values have been adopted from Condon and Shortley [3] and from Garcia and Mack [4] respectively. The effect of the hyperfine structure is almost negligible for our spectroscopic purpose, and has been taken into account with respect to the level $2S_{1/2}$ alone in our calculation. A modification of F, in which the reference wave number is equal to the wave number of the strongest component $2P_{3/2}-3D_{5/2}$, ν_1, instead of ν_0, is illustrated in figure 1. As usual, the x- and y-axes correspond to the real and the imaginary parts of this modified Fourier transform.

The observational data including the theoretically estimated correction $-\pi^{-1}\arg F$, $2\nu_0 D+\pi^{-1}\arg F'-\pi^{-1}\arg F$, are shown in the second column of table 1. Deviation from proportionality to the path difference of these order numbers is still quite significant after this correction, which means a considerable amount of intensity anomaly in either of the fine-structure components. A preliminary survey reveals that the intensity anomaly amounts to about 8 percent of the total intensity and can be ascribed to a suppression in the "red" group or an enhancement in the "blue" group. On the other land, it is highly unlikely that an anomaly by as much as 8 percent of the total intensity would be assigned to either of the five weaker compounds, for which the maximum intensity share is only 11 percent. We have therefore to consider only two cases: suppression in $2P_{3/2}-3D_{5/2}$ or enhancement in $2P_{1/2}-3D_{3/2}$. The former assumption is found to be much more consistent with the experimental data listed in table 1 than the latter. We thus finally assign the anomaly to $2P_{3/2}-3D_{5/2}$ alone.

The 18 values of $2\nu_0 D+\pi^{-1}\arg F'-\pi^{-1}\arg F$ shown in the second column of table 1 now permit least-squares determination of the two parameters, the apparent Rydberg constant at $j=10$ mA, $R_{H,10}$, and the ratio of decrease in intensity of $2P_{3/2}-3D_{5/2}$ to the theoretical total intensity, r, since ν_0 corresponds to $R_{H,10}(1/2^2-1/3^2)$ and since F' can be regarded as a modification of F, in which the intensity distribution of $2P_{3/2}-3D_{5/2}$ is equal to $f(\nu_1)-r\sum_k f(\nu_k)$ instead of $f(\nu_1)$. The results are

$$R_{H,10}=109\,677.604\,5\pm0.003\,2\ \text{cm}^{-1} \qquad (1)$$

and

$$r=7.85\pm0.22\ \text{percent}.$$

The residual errors of the respective values in the second column are indicated in the third column.

3. Correction to Zero Field

In order to reduce the value in eq (1) to the value corresponding to the unperturbed hydrogen atom, it is necessary to determine $R_{H,j}$ at various values of j's and make an extrapolation to $j=0$. Meanwhile, the variation in order of interference as a function of j at an arbitrary path difference is in general affected not only by electrodynamic level shifts common to all the fine-structure components but also by change in the parameter r. Increase in r by a certain amount, however, corresponds to a simple displacement of the origin of coordinates of figure 1 toward the positive direction of the x-axis by the same amount, so that the intersections of the curve with the x-axis give the values of singular path difference where the effect of r disappears. We find two such values, 1.60 and 3.24 cm, but the former is quite inadequate for our purpose because of the extremely low visibility of interferograms.

TABLE 2. *Observed variation in $R_{H,j}$.*

Pair of j's	Variation in $R_{H,j}$	Residual error
mA	$\times10^{-4}$ cm^{-1}	$\times10^{-4}$ cm^{-1}
10, 5	130	+31
	107	+8
11, 6	48	−87
	108	−27
12, 7	233	+54
	219	+40
13, 8	233	+3
	189	−41
14, 9	286	−4
	369	+79
15, 10	308	−51
	323	−36
11, 5	102	−41
	163	+20
12, 6	227	+36
	189	−2
13, 7	228	−20
	213	−35
14, 8	362	+47
	281	−34
15, 9	446	+53
	367	−26

At the path difference of 3.24 cm, the variation in order of interference has been determined twice at each of the eleven pairs of j's indicated in the first column of table 2. The observational data divided by $(1/2^2-1/3^2)\times 3.24$ cm and reduced to variation in $R_{H,j}$ are shown in the second column.

All the results appear as "blue" shift with increasing current and moreover it is a general tendency that the blue shifts appear larger at pairs of larger currents. It is therefore very likely that $R_{H,j}$ will be a monotonically increasing and convex function of j for values of j between 5 and 15 mA, and that this relation will also hold for j lower than 5 mA. We thus assume a simplest empirical relation

$$R_{H,j}-R_H=a(j/10\text{ mA})^b, \tag{2}$$

where a and b are the parameters to be determined by the method of least squares from the 22 values shown in the second column of table 2. The results are

$$a=0.010\,8\pm 0.001\,4\text{ cm}^{-1} \tag{3}$$

and

$$b=3.59\pm 0.30.$$

The residual errors of the respective values in the second column are indicated in the third column.

4. Conclusion

We obtain, from eq (2),

$$R_H=R_{H,10}-a$$

and therefore, from eqs (1) and (3),

$$R_H=109\,677.593\,7\pm 0.003\,5\text{ cm}^{-1}, \tag{4}$$

where the assigned uncertainty is the one-standard-deviation error obtained from the least squares calculations.

No other sources of bias seems to be present except the two assumptions introduced in the course of our reduction of data: omission of intensity anomalies except $2P_{3/2}-3D_{5/2}$ and assumption of eq (2). Tentative calculations under different assumptions, for example intensity anomalies in some other components or some other representations of $R_{H,j}$ in place of eq (2), lead to results slightly different from eq (4), but the difference does not exceed the assigned standard deviation in any case.

5. Comparison With Other Results

We have so far been aware of four independent interferometric results on the determination of the Rydberg constant. These results, reported by Houston [5], by Chu [6], by Drinkwater, Richardson, and Williams [7], by Csillag [8], and recompiled by Taylor, Parker, and Langenberg [9], are unfortunately discrepant by as much as 0.02 cm^{-1} with our result. We also have heard of two other works in progress in the same field of research from Kessler and Roesler [10] and from Stoner [11].

6. References

[1] Steel, W. H., Optica Acta **11**, 211 (1964).
[2] Cohen, E. R., Phys. Rev. **88**, 353 (1952).
[3] Condon, E. U., and Shortley, G. H., The Theory of Atomic Spectra (Cambridge Univ. Press, London), p. 134 (1957).
[4] Garcia, J. D., and Mack, J. E., J. Opt. Soc. Am. **55**, 654 (1965).
[5] Houston, W. V., Phys. Rev. **30**, 608 (1927).
[6] Chu, D. Y., Phys. Rev. **55**, 175 (1939).
[7] Drinkwater, J. W., Richardson, O., and Williams, W. E., Proc. Roy. Soc. (London) **A174**, 164 (1940).
[8] Csillag, L., Acta Phys. Acad. Sci. Hung. **24**, 1 (1968).
[9] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[10] Kessler, Jr., E., and Roesler, F. L., these Proceedings.
[11] Stoner, Jr., J. O., private communications (1970).

DISCUSSION

A. H. Cook: Well, it's certainly very good to have a new measurement of the Rydberg constant and one that does not show a monotonic trend with time.

P. Giacomo: What about the pressure effect?

T. Masui: Oh, it was not possible to observe the pressure dependence of the wavelength since the lamp is sealed off like this (*indicating*). But I think it's necessary and also sufficient to make the correction to zero field by dependence on current intensity.

E. R. Cohen: Do you care to comment on the comparison of your result with Csillag's since the difference is outside the standard errors that are quoted?

T. Masui: Yes. Let me compute chi square for the set of five existing values. It is equal to—I don't remember exactly—but very nearly equal to 8. And this value for a degree of freedom of 4 is not improbable by 8 percent.

Determination of the Rydberg Constant from He II Line Measurements

E. G. Kessler, Jr.

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

and

F. L. Roesler

University of Wisconsin, Madison, Wis. 53706

Detailed studies of the relative positions, relative intensities, and line widths of the He II transitions at 4686, 6560, and 10124 Å excited in a cooled hollow cathode have been made with a pressure-scanned Fabry-Perot spectrometer. These studies indicate that absolute wavelengths of some of the fine-structure components can be measured with sufficient accuracy to improve the precision of the Rydberg constant. A description of the absolute wavelength experiment is given. Large He II line widths limit the accuracy of the absolute wavelength measurements. Two experiments involving excitation processes which show promise of producing narrower lines are briefly described.

Key words: Excitation processes; fine structure; helium; Rydberg; wavelength.

In the most recent adjustment of the fundamental physical constants by Taylor, Parker, and Langenberg [1], the Rydberg constant for infinite mass was assigned the value of 109 737.312±0.011 cm^{-1} on the basis of spectroscopic data on H, D, and He^+ obtained by Houston (1927) [2], Chu (1939) [3], Drinkwater, Richardson, and Williams (DRW) (1940) [4], and Csillag (1968) [5]. In each of these experiments the absolute wavelengths of blends involving two, three, or four components were measured. These measurements were corrected for blending using several assumptions in order to arrive at the most accurate value for the Rydberg constant.

The precision of the Rydberg constant is limited by two sources of uncertainty. The first is the accuracy with which the wavenumbers of the blended peaks of H, D, and He II can be determined. Houston, Chu and DRW quote errors of 2 to 12 mK (mK = 10^{-3} cm^{-1}) with an average uncertainty of 5 mK, while Csillag gives errors of 2 to 3 mK. The second source of error lies in the uncertain assumptions used in adjusting for blending. To derive the Rydberg from the wavelength measurements, relative intensities must be assigned to the components comprising a blend. Theoretical statistical relative intensities were assumed in the above mentioned experiments. However, the work of Roesler and Mack (1964) [6] and DRW indicates that theoretical intensities may not have been present. There is also uncertainty in deciding which components comprise a given blend, since the line profiles and widths are not accurately known. This makes it difficult to judge how far a component must be removed from the peak of a blend before it contributes negligible shift to the peak position.

In the past decade the ionized helium transitions at 4686 Å $n=3$–4, 6560 Å $n=4$–6, and 10 124 Å $n=4$–5 have been studied in great detail by Roesler and Mack [6], Larson and Stanley [7], Berry and Roesler [8], and Roesler and Kessler [9]. The resolution achieved in these experiments far surpasses that obtained by Houston and Chu. The new experiments employed pressure scanned multiple Fabry-Perot spectrometers with photoelectric detection to resolve the complicated structures. The ionized-helium spectrum was excited in a hollow-cathode discharge in all these experiments except that of Larson and Stanley, in which an atomic beam source was employed. The improved resolution along with the relative intensity, relative position, and line-profile information should allow absolute wavelength measurements of He II to be made with sufficient accuracy to improve the value of the Rydberg constant.

A representative spectrum of the 4686 Å line excited in liquid-nitrogen-cooled aluminum hollow cathode is shown in figure 1. The transition consists of thirteen components the more intense of which are drawn on the figure. The resolution which is limited by the width of the helium lines can be slightly improved by employing liquid helium cooling. Berry and Roesler have observed only a 20 percent decrease in line width when liquid helium was employed as a coolant in place of liquid nitrogen. They have further shown that the large line width results from the transfer of momentum during the excitation process. Because the resolution is not

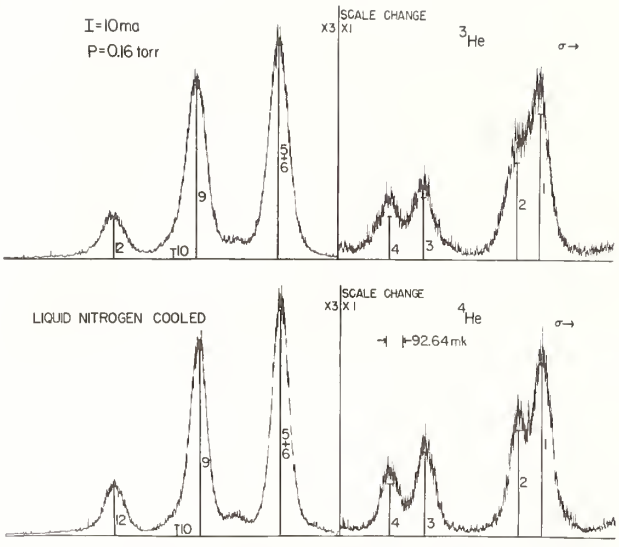


FIGURE 1. *The* He II 4686 Å *spectrum excited in a liquid nitrogen-cooled aluminum hollow cathode.*
The top scan is for ^{3}He and the bottom scan is for ^{4}He. Note the intensity scale change of 3:1 in each scan; the components on the left have been reduced by a factor of three.

greatly improved by the use of liquid helium cooling, we plan to use liquid nitrogen cooling throughout our experiment. Figure 2 shows the dependence of relative intensities on pressure. Components 3 and 12 which arise from upper S states are most sensitive to pressure. This figure clearly demonstrates that intensities other than theoretical relative intensities may be present.

Spectra such as these are decomposed with the aid of a computer. An experimental or numerical profile is determined and the computer is instructed to adjust component positions, intensities, and widths in an iterative fashion to obtain the best least-squares fit. Table 1 gives the positions of the components relative to component 9. Components 1, 2, 3, 4, 5 and 6, 9 and 12 are sufficiently well resolved that their relative positions can be determined very accurately. To estimate the accuracy

TABLE 1. *Fine-structure positions for the* He II
$n=3$–4 line complex relative to component 9
($3D_{5/2}$–$4F_{7/2}$) expressed in mK (10^{-3} cm^{-1}).

Component	^{4}He[a]	^{3}He[a]	Theory[b]
1	1944.4±0.9	1946.4±0.7	1945.31
2	1806.8±1.7	1811.9±2.3	1807.07
3	1273.4±1.1	1276.6±2.6	1273.47
4	1075.2±0.9	1070.2±1.1	1074.96
5+6	454.4±0.4	454.7±0.7	455.53
9	0.0±0.5	0.0±0.8	0.00
10	−113.9±5.3	−114.8±7.0	−122.01
12	−462.9±2.0	−464.1±2.1	−461.92

[a] Uncertainties are one standard deviation.
[b] The theoretical positions were calculated from Garcia and Mack [10]. For the blended components the theoretical position is the center of gravity of the blend based on statistical relative intensities.

TABLE 2. *Fine structure positions for the* He II
$n=4$–6 line complex relative to component 5+6
($4D_{3/2}$–$6F_{5/2}$, $4P_{3/2}$–$6D_{5/2}$) expressed in mK (10^{-3} cm^{-1}).

Component	^{4}He[a]	^{3}He[a]	Theory[b]
1	660.1±2.4	663.4±3.8	659.13
2	601.6±2.7	606.4±3.9	600.56
3	455.9±8.8	475.9±10.0	460.10
4	376.3±4.0	387.0±4.1	383.64
5+6	0.0±0.5	0.0±0.6	0.00
9+10	−207.4±1.2	−208.0±1.5	−207.05
15	−306.3±1.6	−308.4±2.2	−307.37

[a] Uncertainties are one standard deviation.
[b] The theoretical positions were calculated from Garcia and Mack [10]. For the blended components the theoretical position is the center of gravity of the blend based on statistical relative intensities.

with which absolute wavelengths can be determined it is necessary to increase the error bars by at least the uncertainty in determining the exact order numbers. If order numbers are determined to 1/1000 of an order, we will introduce errors of appoximately 1 mK. If we increase the error bars by 2 mK we have a reliable estimate of the accuracy with which we can determine absolute wavenumbers.

A representative spectrum of the 6560 Å line excited in a liquid nitrogen-cooled aluminum hollow cathode is shown in figure 3. This transition consists of 19 components but only the more intense components are drawn on the figure. The relative positions of these more intense components are given in table 2. The relative positions of components 1, 2, 5 and 6, 9 and 10, and 15 can be measured with sufficient precision so that their absolute positions should provide useful information in determining the Rydberg constant.

The 10 124 Å line also excited in a liquid-nitrogen-cooled aluminum hollow cathode is shown in figure 4. This line also consists of 19 components, but only components 1, 2, 3, 4, 5 and 6, 9 and 10, 13 and 18 are drawn on the figure. The accuracy with which the relative positions were measured is shown in table 3. Absolute wavelength measurements on com-

TABLE 3. *Fine structure positions for the* He II
$n=4$–5 line complex relative to component 5+6
($4D_{3/2}$–$5F_{5/2}$, $4P_{3/2}$–$5D_{5/2}$) expressed in mK (10^{-3} cm^{-1}).

Component	^{4}He[a]	^{3}He[a]	Theory[b]
1	605.6±2.5	604.9±3.9	606.50
2	546.0±3.2	548.2±4.9	548.21
3	267.6±5.6	266.9±7.8	262.56
4	171.0±4.4	171.8±4.6	173.37
5+6	0.0±0.9	0.0±1.0	0.00
9+10	−178.7±1.4	−181.5±1.8	−180.97
13	−264.0±1.5	−264.4±1.9	−265.39
18	−466.8±3.1	−468.4±4.0	−469.55

[a] Uncertainties are one standard deviation.
[b] The theoretical positions were calculated from Garcia and Mack [10]. For the blended components the theoretical position is the center of gravity of the blend based on statistical relative intensities.

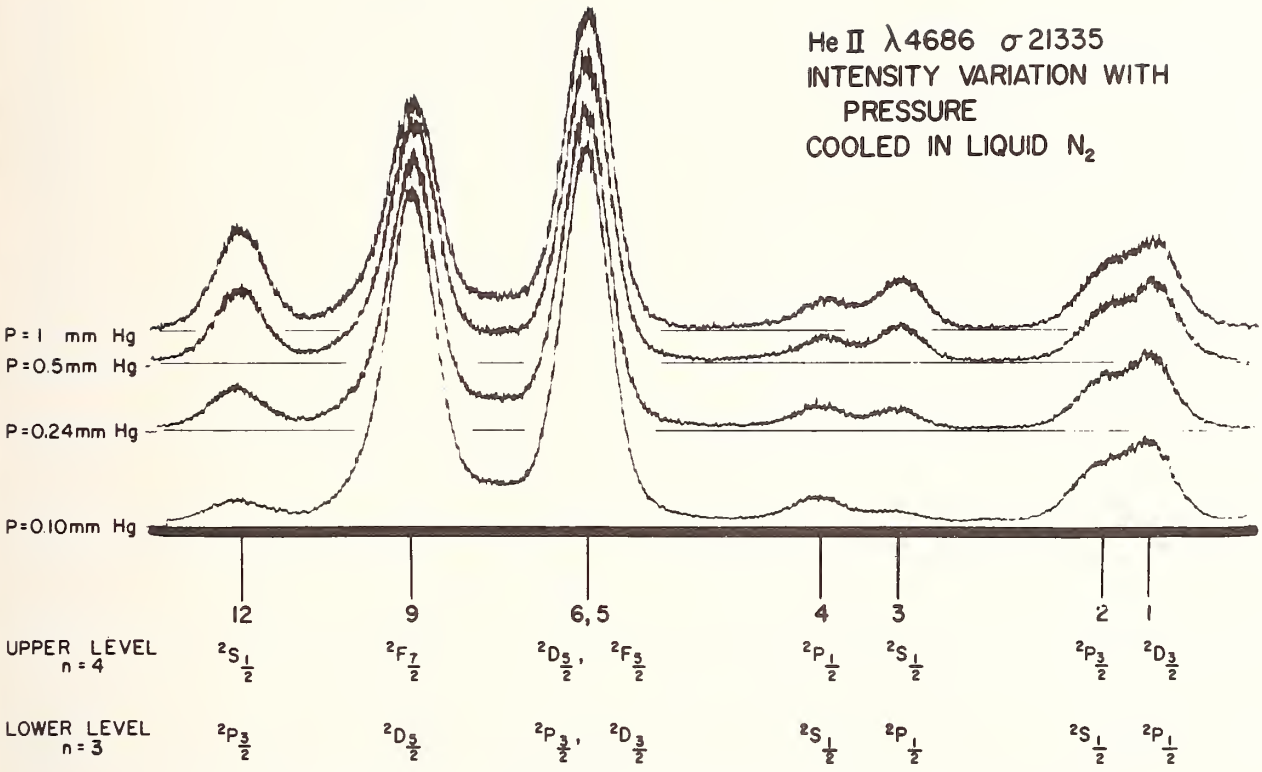


FIGURE 2. *Superposition of traces showing intensity variation with pressure in a liquid nitrogen-cooled hollow cathode plasma.*
The current was 10 mA in each case. The gain for each trace was adjusted to keep the height of the 5+6 blend roughly constant.

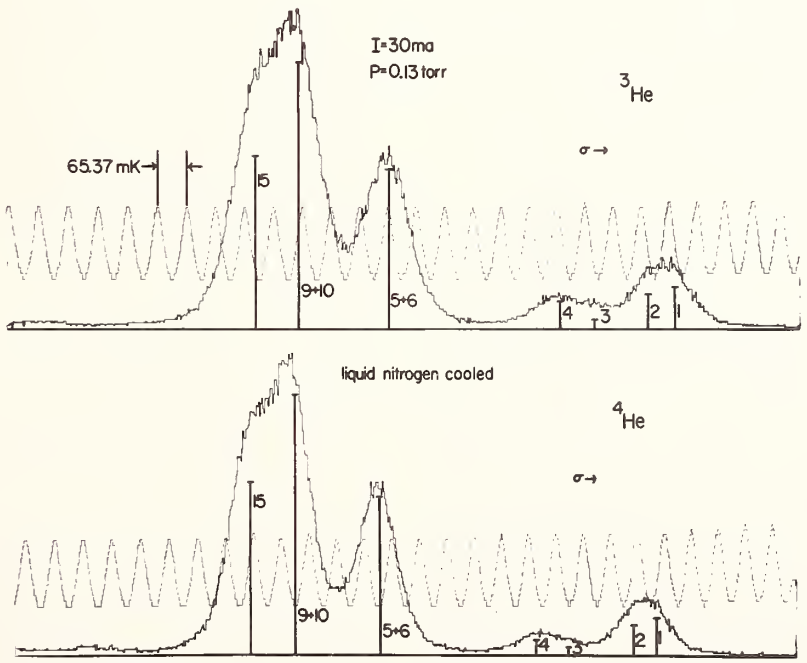


FIGURE 3. *The* He II *6560 Å spectrum excited in a liquid nitrogen-cooled aluminum hollow cathode.*
The top scan is for ^{3}He and the bottom scan for ^{4}He.

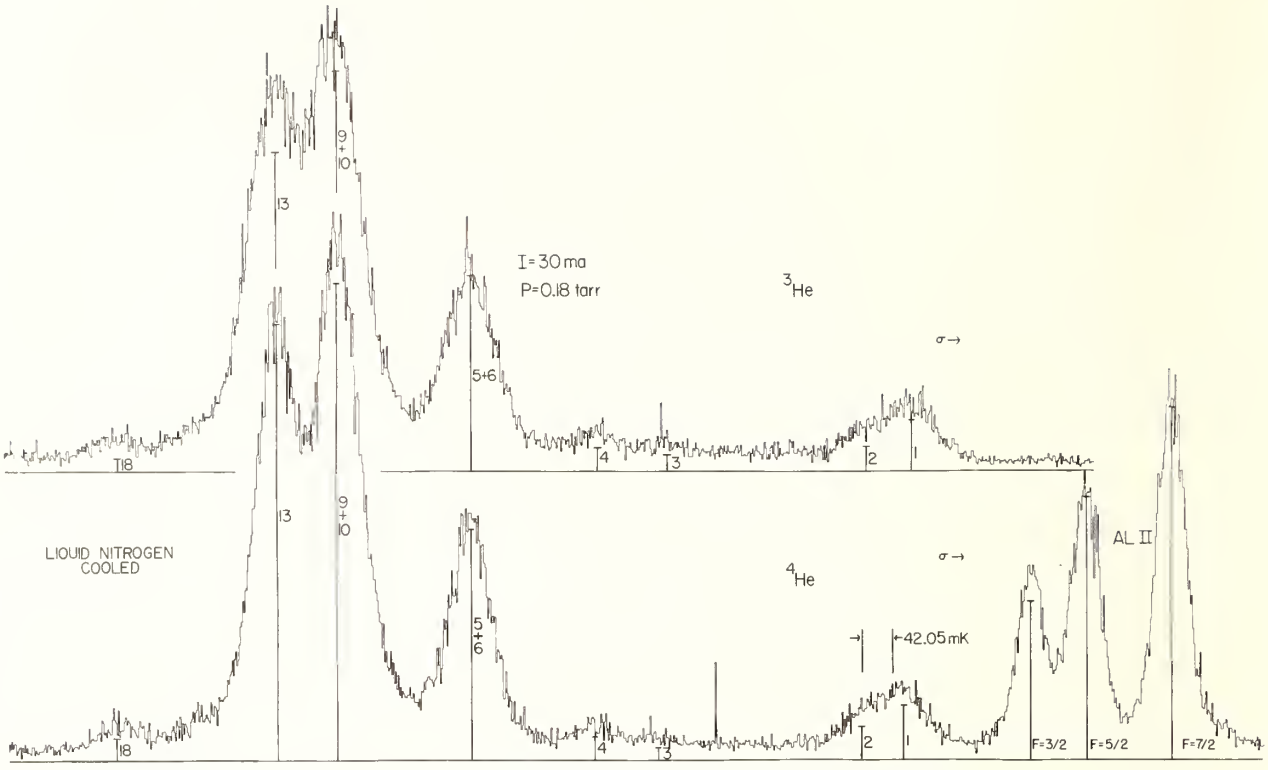


FIGURE 4. *The* He II *10124* Å *spectrum excited in a liquid nitrogen-cooled aluminum hollow cathode.*
The top scan is for ^{3}He and the bottom scan is for ^{4}He. The three lines to the right of the ^{4}He structure are the Al II 3^3D_1–4^3P_0 transition with the three lines arising from the hyperfine splitting of the $^3D_1$ level.

ponents 1, 2, 5 and 6, 9 and 10, and 13 should be precise enough to contribute to the Rydberg determination. Absolute wavelength measurements will be made on 17 components in these three ionized helium transitions, and from them the Rydberg constant will be determined. Previously only components 5 plus 6, and 9 in the 4686 Å transition have been used in deriving the Rydberg from ionized helium data.

The optical arrangement being employed for the absolute wavelength measurements is shown schematically in figure 5. The ionized helium spectrum is excited in a cooled double-anode hollow cathode to allow correction for the shift in wavelength of the fine structure components of He II when the hollow cathode is viewed in two directions, 0° and 180° to the anode–cathode axis [11]. This shift results from the drift of the positively charged helium ions toward the cathode. It has been investigated thoroughly by Berry and Roesler [8] who found that for cathodes larger than 5/8 inches in diameter the drift shift was smaller than the uncertainty in the position measurements at the optimum discharge conditions for Rydberg measurements. Accordingly, we have minimized the drift shift by using cathodes having diameters larger than 5/8 in. Nevertheless, scans are alternately recorded with the ions drifting toward and away from the spectrometer to allow detection of and correction for any residual drift shift.

The hollow cathode radiation is analyzed by a Fabry-Perot spectrometer similar to ones thoroughly described in the literature [12–14]. The light passes through a grating premonochromator operated at low resolution, through two Fabry-Perot étalons in separate pressure chambers, and is focused onto an aperture A (figure 5). A photomultiplier is used to detect the light passed by the spectrometer. A double étalon system is used so that overlapping orders are suppressed. However, because there are no repeated orders an auxiliary calibration system must be

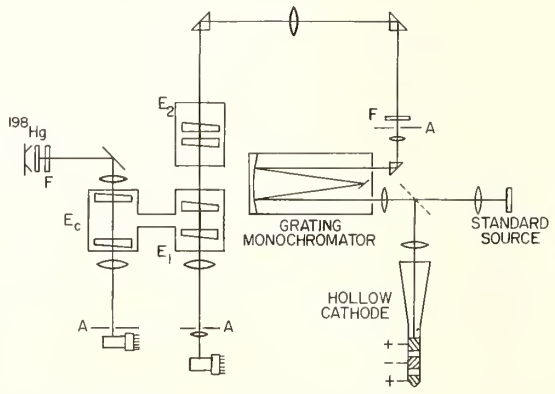


FIGURE 5. *Sources and spectrometer.*
E_1, E_2, and E_c are the resolving étalon, suppression étalon, and the calibration étalon respectively. F are narrow pass filters, and A are circular apertures.

employed to provide a wavenumber scale for the spectrum. A third long spaced étalon illuminated by a water cooled ^{198}Hg electrodeless discharge source and scanned in unison with the main étalon provides precise interferometric calibration of the spectrum. The calibration fringes are detected with a second photomultiplier and are recorded simultaneously with the ionized helium spectrum.

The procedure for taking measurements is as follows: First the spectrum is taken with the double etalon to obtain relative positions, line widths, relative intensities, and line profiles. Subsequently etalon E_2 is removed from the light beam and the ionized helium spectrum and several standard wavelengths from a water cooled ^{198}Hg electrodeless lamp are alternately passed through étalon E_1 and recorded. The method of exact fractions is employed to determine the exact wavelengths. Although there will be overlapping of orders in the helium spectra, our knowledge of relative intensities, line widths, relative positions, and the line profile will enable us to decompose the spectra to obtain accurate measurements of the component positions. Several different spacers will be used to avoid phase shift problems and to change the overlapping pattern.

The technique of obtaining absolute wavelength measurements with a scanning interferometer requires special consideration. In the usual photographic procedure the interferometer is illuminated with the spectrum of interest and the standard line simultaneously, and an auxiliary spectrograph is used to separate the Fabry-Perot patterns. Any discrepancies due to slight changes in temperature or pressure in the étalon are thus avoided. Photoelectric detection provides linearity of response and relatively high quantum efficiency, but requires sampling small regions of the spectrum successively. Thus if one chooses to employ photomultipliers, one must scan the spectrum by some means such as varying the pressure in the étalon cavity. In addition the standard line and spectrum of interest must be viewed alternately to avoid superposition of the standard line on the spectrum of interest. In the application of the method of exact fractions, all fractional orders must be measured at the same optical spacing. The long-spaced auxiliary étalon now serves as an indicator to assure that the fractional orders are measured at identical optical spacings. Temperature changes in the same étalon between successive scans and temperature differences between the two étalons in the same scan can cause order shifts much larger than the accuracy with which we can measure fractional orders. The increases and decreases of the gas pressure inherent in pressure scanning can result in changes in the gas temperature of several degrees unless some precautions are taken.

In order to keep the pressure chambers, étalon holders and plates, and scanning gas at a constant temperature, we have coupled the étalon holders and plates to the pressure chambers and the large étalon cavity to the étalon holder by sets of copper fins or slats. For narrow spaced étalons (<10 mm) there is sufficient coupling between the cavity gas and the holder and plates so that fins are not necessary. The pressure chambers are kept at constant temperature by means of circulating water. With such a system temperatures have been kept constant to within a 0.01 °C, implying shifts of 0.0002 orders which is below our measuring capability.

The precision of the measuring technique is many times higher than the precision with which the line centers can be presently determined due to the width of the components. The studies of Berry and Roesler [8] and of Larson and Stanley [7] have shown that the lower limit of the width of the lines obtained by direct electron excitation is fixed by momentum transferred from the exciting electron in the process $He+e\rightarrow He^+(n,\ l)+2e$, which is dominant in the atomic beam and the hollow cathode. Consequently, negligible improvement in precision over that achieved with a liquid nitrogen-cooled hollow cathode can be obtained by liquid-helium cooling the discharge or with the electron-excited atomic beam.

Two different excitation processes now being carefully studied show some promise of yielding significantly narrower lines and could eventually be used to achieve higher precision than is presently foreseen for the Rydberg measurements. The first is the photoionization-excitation process $He+h\nu\rightarrow He^+(n, l)+e$. Photons with wavelengths shorter than 164 A can excite the $n=4$ levels. The emission rate of 4686 Å radiation excited by this process using the XUV flux from the 250 MeV electron storage ring at the Physical Sciences Laboratory at the University of Wisconsin has been measured at low spectroscopic resolution, and is sufficient to make it appear attractive for eventual Rydberg measurements. High-resolution measurements with liquid helium-cooled helium gas are scheduled for late 1970. The momentum transfer at threshold is negligible. For photon energies above threshold there is excess energy to be shared between the ejected electron and the excited ion which will contribute appreciable Doppler broadening. Nevertheless, by appropriate prefiltering of the XUV flux to cut out the highest energy photons we anticipate line width as low as 27 mK, approximately four times smaller than achieved with the liquid nitrogen-cooled hollow cathode.

A second process that appears attractive is three body recombination, $He^{+\ +}+2e\rightarrow He^+(n, l)+e$, which dominates the one-electron recombination process except at very low electron densities [15]. The $He^{+\ +}$ ions are formed by a 1 μs current pulse in cooled helium gas at a pressure around 0.3 torr and a temperature of 77 K. The ions are allowed to cool for approximately 10 μs, during which time all directly excited $He^+(4, l)$ states decay, and then the detector is switched on for several microseconds. The signal strengths and line widths appear attractive for consideration of this process for Rydberg measurements, and moreover components 3 and 12 of the 4686 A transition vanish because of the slow rate of population of the 4s level in this process. At the time of writing, high-resolution experiments with a liquid-helium-cooled pulsed discharge are being carried out.

Acknowledgments

The authors take pleasure in acknowledging the work of D. Tracy on the photoionization-excitation process and of W. Kuhlow on the recombination process. Parts of this work have been supported by the National Science Foundation.

References

[1] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[2] Houston, W. V., Phys. Rev. **30**, 608 (1927).
[3] Chu, D. Y., Phys. Rev. **55**, 175 (1939).
[4] Drinkwater, J. W., Richardson, O., and Williams, W. E., Proc. Roy. Soc. (London) **174**, 164 (1940).
[5] Csillag, L., Acta Phys. Acad. Sci. Hung. **24**, 1 (1968).
[6] Roesler, F. L., and Mack, J. E., Phys. Rev. **135**, 58 (1964).
[7] Larson, H. P., and Stanley, R. W., J. Opt. Soc. Am. **57**, 1439 (1967).
[8] Berry, H. G., and Roesler, F. L., Phys. Rev. **1A**, 1504 (1970).
[9] Kessler, Jr., E. G., Ph. D. Thesis, University of Wisconsin, 1969.
[10] Garcia, J. D., and Mack, J. E., J. Opt. Soc. Am. **55**, 654 (1965).
[11] Roesler, F. L., and DeNoyer, L. K., Phys. Rev. Letters **12**, 396 (1964).
[12] Chabbal, R., J. Rech. Centre Natl. Rech. Sci. Lab. Bellevue (Paris) **24**, 138 (1953); (English transl.: Document AERE Lib./Trans. 778, Harwell, England).
[13] Chabbal, R., Rev. Optique **37**, 49 (1958).
[14] Chabbal, R., and Jacquinot, P., Rev. Optique **40**, 157 (1961).
[15] Hinnov, E., and Herzberg, J. G., Phys. Rev. **125**, 795 (1962).

DISCUSSION

G. W. SERIES: I should like to make a statement and ask whether you would agree. It seems to me that however one excites the line one wishes to study, one has to make up one's mind whether one is interested in studying the fine structure and in checking that or in measuring the Rydberg. If one is satisfied with the theory of the fine structure, then one should concentrate all one's efforts on measuring the strongest, cleanest components one can find in the structure, even if there is only one component.

E. KESSLER, JR.: I agree with that statement, but I would like to qualify my agreement. It is very difficult to obtain one clean component. As a result it is necessary to worry about the minor components which overlap the stronger components. I agree that one should concentrate on the stronger components. If you can measure one clean, strong component, that's probably better than 15 that overlap.

G. W. SERIES: Could I make a supplementary remark which really is to make a point I omitted from my own talk through shortage of time.

In your He^+ line at 4686 Å, I think there are possibly two components that are strong and clean.

E. KESSLER, JR.: You're talking about components 9 and 5 plus 6?

G. W. SERIES: I think so.

E. Kessler, JR.: These are the two strongest components.

G. W. SERIES: Moreover, they are not very sensitive to Stark effect.

E. KESSLER, JR.: That's exactly right.

G. W. SERIES: In the Balmer alpha also there is a strong component with a satellite adjacent to it that one could probably not resolve, but the important thing is that these are both part of the same fine structure doublet. It's from the $3D_{5/2}$ and $3D_{3/2}$ levels I think, so that one could apply some theory quite reliably in that case.

E. Kessler, JR.: I see what you're saying.

G. W. SERIES: So I think in the 4 and 3 transition of He^+ and also in Balmer alpha one has these conditions satisfied.

E. KESSLER, JR.: Yes. That's correct.

P. GIACOMO: May I ask whether both your étalons are swept in pressure?

E. KESSLER, JR.: Yes, both étalons are swept in pressure, and a pressure difference is maintained between the étalons.

P. GIACOMO: Then I think that the error on fine structure is eliminated, but it may be a source of error when measuring different wavelengths because the matching of the two étalons may be different for the two wavelengths.

E. KESSLER, JR.: I agree that it would be necessary to retune a double étalon system when you scan a different wavelength. However, with a single etalon system there is no tuning.

A. JAVAN: I am somehow curious. All along in these talks the Stark effect in the discharge has been mentioned a few times. Could someone comment on roughly how much is the order of correction that you have to make on the spectrum that you see? How large is this effect in a discharge usually? Is it a part in 10^6 or 10^7 or it is close to the edge of the error?

E. KESSLER, JR.: I think that you will find that it is smaller than a part in 10^8. We know that the fields are a few volts per centimeter from measurements of the drift shift. With fields of this magnitude shifts are in the order of tenths of milliKaisers or smaller.

A. JAVAN: Well, is it the same thing with atomic hydrogen?

G. W. SERIES: If I could also answer the question: I would have thought that one can so monitor the Stark effect that one doesn't have to apply any correction at all on the component that one is interested to measure for the Rydberg because one is monitoring on different components, more sensitive ones.

E. KESSLER, JR.: Might I say that in the 4686 Å transition the components 3 and 4 are most sensitive to Stark effect. If there is any Stark effect these components will split apart and their separation will be noticeably larger. The components which one would probably use, the 5 plus 6 and the 9, will be much less sensitive to the Stark shift.

ELECTRICAL STANDARDS

From over the world did they roam,
Trading thoughts contained in the dome.
The Ampere and Volt
Were given a jolt;
Then most of them quickly went ohm.

ANON.

The Fundamental Electrical Standards: Present Status and Prospects for Improvement

Robert D. Cutkosky

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Absolute measurements of resistance and current have until recently been based on instruments involving carefully constructed and measured single layer solenoids. Minor improvements in technique have resulted in a continuous reduction of measurement uncertainty, but overall uncertainties much smaller than one part in 10^6 are not likely to be achieved without radical changes in design. Some of the principal sources of uncertainty in these experiments are analyzed, and, when possible, alternative measurement systems are proposed which could lead to higher accuracy. Within the time scale appropriate for measurements of this kind, absolute resistance measurements based on calculable capacitors are quite recent, even though the first such measurement was reported in 1961. More recent calculable capacitors have resulted in absolute resistance measurements of higher accuracy than may reasonably be expected from systems utilizing calculable solenoids. The possibility for improvements in absolute current or absolute voltage measurements does not seem quite so clear, but some ideas are presented which are either now being tried or are worth some consideration. An attempt is made to indicate the more important sources of uncertainty in the proposed systems.

Key words: Capacitors; current balances; electrical measurements; electrical standards; electrometers; inductors.

1. Introduction

It is unfortunate that the completion of experimental determinations of physical quantities and of measurements related to them do not always coincide with the convening of an international conference of appropriate stature. In the case of electrical standards, many years commonly elapse between the inception of a particular piece of work and its completion, and in view of the fact that only a few laboratories in the world are involved in such measurements, it may not be surprising that this paper contains no results.

In the short space available I shall quickly summarize the more recent absolute measurements of capacitance, resistance, voltage and current, and give the most attention to the present status of work now underway and to various proposals for improved measurement systems. Some comments on the possible future use of Josephson junctions as maintenance standards of voltage will also be made, although physically the Josephson effect is most useful as a means of accurately measuring e/h.

The starting point for all absolute electrical measurements rests in the extensive body of classical electromagnetic theory, by means of which one can in principle calculate the magnetic fields, inductances, and forces associated with current-carrying conductors, and the electric fields, capacitances, and forces associated with charged conductors, from mechanical measurements of the shapes and positions of these conductors. The task of the metrologist involved in this work is to determine a suitable geometry for his equipment so that the necessary mechanical measurements can be made with a high accuracy, and so that the resultant electrical property has such a magnitude that it can be accurately compared with a set of fixed "maintenance" standards which preserve the unit so obtained. Many compromises must be made in the design of absolute electrical measurement equipment to minimize the uncertainty of the final result. These compromises involve among other things considerations of the relative ease with which various systems can be fabricated and measured.

2. Units of Impedance

Two quite distinct approaches to an absolute ohm measurement are available. One approach begins with an inductor whose inductance can be calculated from its mechanical dimensions, and the other begins with a capacitor whose capacitance can be calculated from its mechanical dimensions. The second method requires a knowledge of the speed of light, which is generally treated as an auxiliary constant.[1] In principle either a calculable inductor or a calculable capacitor can be used to assign a value to a standard resistor using purely bridge methods, and requiring only a measurement of frequency in addition to suitable impedance ratio measurement techniques. Generally these impedance comparisons can be made with substantially better accuracy than is presently achievable in the construction and evaluation of either a calculable inductor or a calculable capacitor.

2.1. The Ohm from Inductance

Calculable self inductors are generally constructed by preparing a grooved cylindrical former of a stable material such as fused silica, and placing a single-layer helical winding of copper wire in the groove. The dimensions of the completed solenoid are measured as precisely as possible and from these measurements the self inductance of the solenoid can be calculated.

Calculable mutual inductors usually contain one precision solenoid similar to that described above, but with one or more gaps in the helical winding, placed so that a region exists exterior to the solenoid in which the magnetic field is very small. The secondary winding is placed in this low-field region, since it is then not necessary for its dimensions or its location to be measured with a high accuracy. Multilayer windings can be used for the secondary without seriously affecting the accuracy with which the mutual inductance can be calculated. Alternatively, a mutual inductor can be constructed by placing a bifilar winding on a single cylindrical former.

At least six absolute ohm determinations based on calculable inductors were reported since 1955. Most of these measurements utilized mutual inductors rather than self inductors, partly because self inductors have larger phase angles than mutual inductors, which complicates the electrical part of the measurement. Two of the recent inductor-based absolute ohm determinations were assigned uncertainties of about 2 parts in 10^6 [1, 2]. In both cases, the uncertainties were almost entirely attributable to the difficulties involved in measuring the mechanical dimensions of the single-layer, wire wound solenoids.

Some reduction in the mechanical measurement uncertainties for solenoids may be possible, but it seems likely that an uncertainty substantially smaller than 1 part in 10^6 will not be achievable with present inductor designs. The reasons for this are first, that even single layer solenoids wound on rigid, grooved formers are slightly compressible, and uncertainties remain even after corrections are applied for the distortion caused by the pressure of the micrometer used in the measurement of solenoid diameter and pitch. Second, oxide films on the wire cause formidable problems; and third, the distortion of the wire caused by winding it tightly on the former cannot be easily determined. Finally, it is necessary to estimate the current distribution in the wire. This depends upon how the resistivity of the wire varies with distance from the axis of the solenoid, and is a function of wire diameter, solenoid diameter, winding tension, and the material chosen for the wire, which is usually copper. Eddy currents within the wires also affect the current distribution,

[1] The vacuum capacitance of a capacitor is equal to a geometric factor times the electric constant ϵ_0, where $\epsilon_0\mu_0=1/c^2$. The magnetic constant $\mu_0=4\pi\times10^{-7}$ henry/meter in the SI system.

and are not easy to estimate (eddy currents induced in other metal parts of the system can also cause serious errors if the problem is not recognized).

The current distribution uncertainties discussed above could probably be reduced by reducing the wire diameter; but if carried to an extreme, the wire distortion caused by micrometer pressure would be very large. Additionally, reducing the wire cross section would increase the solenoid resistance, and result in either an increase in power consumption and hence larger temperature uncertainties if the solenoid is in a current carrying part of the electrical circuit, or in increased thermal agitation noise if it is in the potential part of the circuit.

Proposals have been made to produce solenoids by evaporating very thin metal films in a helical pattern on smooth cylindrical forms. There is no doubt that this would result in a reduction of the solenoid diameter uncertainty, but it would be very difficult to determine how the current was distributed over the width of the film. Very small defects in the film could cause significant local variations in the pitch of the solenoid, which could result in substantial errors in the calculated inductance. The relatively large solenoid resistance to be expected with such a system could also be a problem for the reasons cited above. It has been argued that the use of superconducting windings on a solenoid would eliminate many of the above problems, but the current distribution within a superconducting wire is at best ambiguous, and no reasonable way has yet been found for measuring the solenoid dimensions at low temperature.

A completely different geometry for a calculable mutual inductor has been treated theoretically by Page [3]. The proposed system uses a set of eight parallel wires equally spaced around a circle. In this system, the mutual inductances between each pair of wires and every other pair of wires is measured, and these mutual inductances can be related to the length of the wire system. Many of the uncertainties due to small errors in wire placement are of second order, and only the length of the system need be measured to high accuracy. No serious attempt has yet been made to construct such a system, but it seems clear that the problems would be formidable. The mutual inductances are likely to be in the order of 10^{-7} H, which presents serious but not insurmountable electrical measurement problems [4].

2.2. The Ohm from Capacitance

Calculable capacitors have three principal advantages over calculable inductors. Negligible power is lost in a capacitor, so the mechanical dimensions do not change with time due to the resultant change in temperature. The corresponding problem with capacitors has to do with dimensional changes due to the electrostatic forces between the capacitor electrodes, and usually contributes errors smaller than one part in 10^8. The second advantage of capacitors is that the current distribution problem does not exist. In this respect, capacitors are not as pure as might be wished, because dielectric films on the electrode surfaces introduce an analogous uncertainty in the effective electrode spacing. However, the dielectric films are usually quite thin, and the resultant uncertainties are much smaller than the uncertainties resulting from the current distribution in a calculable inductor. The third advantage of calculable capacitors over calculable inductors is that all critical fields are contained within the calculable region, so that the effects of objects external to the calculable instrument can be ignored. This eliminates the need for special nonmagnetic buildings as are generally required for high-precision inductance measurements.

All calculable capacitors now in existence or known to be under construction are of a type called a cross capacitor. This geometry was first proposed by A. M. Thompson and D. G. Lampard, and in practice is usually constructed of four cylindrical electrodes of circular cross section [5]. The axes of the four electrodes are parallel and are located at the corners of a square just slightly larger than the electrode diameter. The two capacitances of interest are the capacitance coefficients between opposite pairs of electrodes. The exterior of the system is completely shielded so that only the region between the four electrodes is of concern.

Provided that reasonable care is taken to obtain the requisite two-dimensional symmetry, all small perturbations from the ideal geometry including those due to dielectric films on the electrode surfaces enter only to second order in the average of the two cross capacitances, leaving only the first order dependence of capacitance upon the overall length of the electrode system. This single critical dimension is measured in the more refined cross capacitors with interferometric techniques.

Several cross capacitors have been described in recent years, but so far only two absolute ohm measurements based on cross capacitors have been completed [6, 7]. The most accurate of these two was the ohm measurement reported by Thompson and making use of the cross capacitor described by Clothier [8]. They reported an overall uncertainty of one part in 10^7 exclusive of the uncertainty in the speed of light. Somewhat smaller uncertainties may be expected soon as improvements are made in the Clothier/Thompson apparatus and as similar systems now under construction elsewhere are completed.

At the present time it appears that capacitor-based absolute ohm measurements are capable of providing much higher accuracy than are inductor-based absolute ohm measurements. The relative accuracies of the two techniques are so disparate that work with calculable inductors could well be abandoned, except that a comparison of measurements made in such completely different ways helps to prevent gross errors. Inasmuch as a capacitor-based absolute ohm measurement requires a knowledge of the speed of light and an inductor-based measurement does not, one can treat the pair of ohm measurements as a speed-of-light measurement, and possibly learn something about the dependence of c upon frequency. No discrepancy attributable to such an effect has

yet been observed, and the continuation of calculable inductor research on this ground might be hard to justify.

3. Units of Voltage or Current

After having established a unit of resistance by means of one of the techniques outlined above, an absolute determination of either the volt or the ampere suffices to establish both units through Ohm's law. In practice it is the unit of voltage that is maintained rather than the unit of current, because of the proven stability of saturated cadmium sulfate standard cells, but in fact all of the precision measurements of this kind made until now have been with current sensitive instruments called current balances. Voltage can also be measured directly with electrometer-type instruments, and some such instruments are now under construction.

From a dimensional standpoint, absolute resistance measurements require only length and time determinations, whereas voltage or current measurements require in addition a determination of mass. The experiments involve a measurement of the torque or force between two current carrying conductors or between two charged electrodes. The electrical force is usually compared with the gravitational force on a known mass, and hence a knowledge of the local acceleration of gravity is also required. These additional complications tend to make absolute voltage or current measurements much more difficult than absolute resistance measurements.

3.1. Current Balances

The most common type of current balance is often ealled a Rayleigh balance, and contains a fixed center-tapped solenoid having equal currents in each half of the solenoid but of opposite direction, to yield a strong radial component of magnetic field in the center of the system. A relatively short current-carrying solenoid is placed in the center of the system and attached to one arm of a precision balance. Reversing the sign of the fixed-solenoid current while keeping the current in the movable solenoid fixed changes the sign of the force on the movable coil. The balance is then restored to equilibrium by adding a known mass to that arm of the balance. Some Rayleigh balances have a set of coils on each arm of the balance, and the movable solenoid (or solenoids) may be either of larger or smaller diameter than the fixed solenoid. Within the last few years, about four Rayleigh-type current balances have been described, with the smallest reported uncertainties being about 4 parts in 10^6 [9], and about 5 parts in 10^6 [10].

One current balance has been described which involves the torque between two current carrying solenoids whose axes are perpendicular [11]. The torque is measured by attaching one solenoid to a balance beam and restoring the balance equilibrium after reversal of the current in the fixed solenoid by placing a known mass on one pan of the balance. The distance between the knife edge supporting this pan and the central knife edge of the balance must also be measured with this system. Some problems were encountered initially with distortion of the balance beam, but this probably does not constitute a fundamental limitation of the method. The most recent ampere determination with this instrument yielded an uncertainty of 7 parts in 10^6 [12].

The solenoids for all types of current balance are normally constructed and measured using the same techniques as are used for calculable inductors, and all of the problems encountered in calculable inductors also occur in current balances, except that since direct current is used in current balances, eddy current effects do not influence the current distribution in the wires. Since the movable solenoid of a current balance must be kept as light as possible to avoid overloading the balance, the movable solenoid is usually small and quite compressible, and its dimensions are therefore difficult to measure accurately. This is a principal contributor to the uncertainty in a current balance experiment.

A second major problem encountered in current balances is caused by air currents generated by the heated wires, which disturb the balance. Presumably no systematic errors are caused by the air currents, but very erratic balance fluctuations are observed, and long measurement sequences are required in order to obtain a reasonably small uncertainty in the measured force or torque.

Until recently the uncertainty in the local acceleration of gravity contributed substantially to the overall uncertainty of an ampere determination. Recent refinements in gravity measurement techniques have made this source of uncertainty negligible.

A substantial reduction in the uncertainty of an absolute ampere determination would seem to require an improved technique for constructing and measuring the movable solenoid. The fixed solenoid can be made large and rigid enough to allow a relatively accurate determination of its dimensions, and schemes exist for comparing the pertinent electrical properties of one solenoid with another. For example, the ratio of the magnetic fields in the centers of two solenoids carrying equal currents is a function of the relative dimensions of the two solenoids, and may be used in conjunction with a set of mechanical measurements of both solenoids to improve the data obtained for the less precisely known solenoid by a least squares process. Only a small improvement in the calculated force constant of a current balance is likely to result unless the two solenoids are specifically designed to take advantage of this technique.

Another approach to a better ampere determination involves placing the movable solenoid in a much stronger field than can be conveniently obtained from a calculable fixed solenoid, and measuring the ratio of this strong field to the much weaker field of a calculable solenoid using proton gyromagnetic resonance techniques. Forces strong enough to be accurately measured can be obtained with this system using a much simpler and more precisely

measurable movable solenoid, and in fact, the movable "solenoid" usually consists of a winding around the edge of a rectangular insulating form. In this case an evaporated metal "winding" might offer some advantages, since the distribution of current over the width of the conductor is of little concern.

A possible flaw in the above system (sometimes called a Cotton balance) is that one must assume that the proton gyromagnetic ratio is independent of field strength. No contrary evidence exists, but lacking further verification, the results of such an experiment might be subject to question. It is interesting to note that the most direct evidence for the independence of the proton gyromagnetic ratio upon field strength is based on measurements conceptually similar to those described above and on measurements with a conventional current balance. Measurements by Huggins and Sanders of the ratio of the proton resonance frequency to the resonance frequencies of ^{3}H, ^{7}Li, ^{11}B, and ^{19}F at both high and low field intensity indicated that the field dependence of the proton resonance frequency is certainly small, but a slight field dependence cannot be ruled out [13].

3.2. Electrometers

The possibility of obtaining the unit of voltage by measuring the force between two charged electrodes has long been intriguing, because of the absence of the air currents which disturb a current balance. Unfortunately the electrometer geometry appropriate for such a system is not usually amenable to the precise mechanical measurements necessary for determining the force constant with adequate precision directly, and this fact has discouraged work in this direction.

One proposed scheme for constructing a directly calculable electrometer utilizes a pool of mercury with a fixed flat horizontal electrode just above the pool [14]. The distance between the electrode and the pool can be measured interferometrically with and without a voltage between the pool and the electrode. The electrical force between the pool and the electrode can then be inferred from measurements of the density of the mercury and of the local acceleration of gravity.

Problems are expected with this system in measuring the mercury density to an adequate accuracy, and in assuring the cleanliness of and lack of surface films on the mercury. The system appears to be feasible, and preliminary investigations have indicated that the mercury surface can be maintained reasonably stable and ripple-free, but the overall uncertainty likely to result cannot yet be estimated.

An indirect approach to an electrometer system makes use of electrodes in the form of interleaving cylinders, for which the electrical force is constant over a substantial axial displacement of one electrode with respect to the other. The force at a given electrode spacing is proportional to the rate of change of capacitance with respect to distance, and since the force is nearly constant, one can show that the average force over a finite electrode displacement is equal to $\frac{1}{2}V^2\,\Delta C/\Delta l$. It is proposed that the total electrode displacement Δl be measured interferometrically, and that ΔC be measured by comparison with a calculable capacitor. The average force can presumably be determined from a large number of discrete force measurements made with different relative electrode displacements.

The above system seems relatively straightforward at first glance, but some difficult problems remain unsolved. For example, if the voltage between the electrode is dc, the capacitance required for use in the equations is the d-c value, and may differ from the more readily measurable ac capacitance obtained in terms of a calculable capacitor using conventional bridge techniques. Slightly lossy dielectric films on the surfaces of the electrodes would produce such a difference, and hence great care would be necessary to obtain clean electrodes. This problem would be avoided if the electrodes were excited with alternating current, but the substantial double frequency force variation that would then result might prove damaging to the knife edges of the balance, and in addition the measured ac voltage would have to be transferred to dc with a thermal transfer instrument for comparison with the voltage of a standard cell.

Any electrometer would be improved if it could be operated in a vacuum, because the buoyancy corrections for the mass and the dielectric constant correction for air would both be eliminated. The mercury-pool electrometer would probably work well under these conditions, although evaporation of the mercury has been reported to cause some problems. Experiments involving the operation of conventional knife-edge balances in a vacuum unfortunately have not been notably successful. It is expected that occluded surface films on the mass standard would be partially released under vacuum conditions, making an assignment of the vacuum mass a rather difficult task. The magnitude of this effect is at the present time not known.

In view of the many difficult problems that one can foresee in the construction and operation of a precision electrometer, it is not possible at this time to estimate the accuracy likely to be obtained. Judging from some preliminary investigations, it seems to be possible to obtain the unit of voltage from an electrometer with an uncertainty at least as small as can now be obtained by means of a current balance, or a few parts in 10^6.

4. Standards for Maintaining Electrical Quantities

There is little point in investing a great deal of effort in an absolute electrical measurement if there are no adequate standards for maintaining the unit, or for use in comparing the measurements made in different laboratories. The currently accepted standards for maintaining the units of voltage and

resistance appear to introduce uncertainties of a few parts in 10^7. Some possibilities exist for improving these standards, but at present only minor improvements can be foreseen.

4.1. Maintenance Standards for Impedance

The unit of resistance has been maintained by means of 1-Ω, dc resistors for at least fifty years. Proposals have been made to use 10^4-Ω resistors for this purpose because of the valid argument that they are not as strongly influenced by thermal voltages in the measurement circuit, and because of the dubious argument that the optimum source impedances of some recently developed galvanometers are closer to 10^4 Ω than to 1 Ω. Improvements in low-impedance galvanometers in the next few years cannot be ruled out, but the uncertainties caused by varying thermal voltages will always be more significant with smaller resistors. If careful attention is given to the construction of 1-Ω resistors and the associated measurement circuit, one might expect to reduce the uncertainties due to thermal voltages to about one part in 10^8, whereas the limit set by thermal agitation noise is closer to one part in 10^9, and is independent of the resistance. Since a calculable capacitor is inherently a high impedance standard, it is somewhat easier to measure a 10^4-Ω resistor in terms of such a standard than it is to measure a 1-Ω resistor. In practice the transfer from ac impedance to dc impedance is usually made at the 1000-Ω level, and as a result only one less measurement is required if the unit is maintained at 10^4 Ω. A procedure for accurately measuring 1-Ω resistors is required in any case, since current balances normally employ currents of about 1 A, which are measured by comparing the voltage produced by this current passing through a 1-Ω resistor with the voltage of a standard cell.

The effects of thermal voltages could be eliminated by maintaining the unit of resistance by means of ac resistors. Existing ac detectors and matching networks are sufficiently well advanced to allow the measurement of any ac resistor between 1 Ω and 10^5 Ω with an uncertainty limited only by thermal agitation noise in the resistor. Capacitors can also be used to maintain a unit of ac impedance, and measurement uncertainties of about one part in 10^9 can in this case be achieved. The absence of self-heating effects in capacitors makes them especially appealing. Existing capacitors appear to drift less than one part in 10^7 per year, but are not quite as stable as the best 1-Ω resistors.

It is probably too early to recommend a change in the mode of maintaining units of impedance. The most practical solution is for laboratories to maintain separate sets of promising standards for several years, and to periodically intercompare the sets with each other. If a particular type of standard is then found to be more suitable than the others, this fact will emerge from a statistical analysis of the data, and no formal abandonment of any type of standard need be made. The maintained unit could be a weighted average of the units maintained by the different types of standards, with the weighting factors changing with time as stability data accumulates. If the uncertainties associated with intercomparing the different sets of standards were very much larger than the uncertainties associated with comparing the standards within a set with each other, it could be useful to recognize more than one maintained unit for a single quantity, but in this case it would be very important to distinguish between the various units.

4.2. Maintenance Standards for Voltage

The unit of voltage is universally maintained by means of saturated Weston cells. These standards are much more stable than any other known electrochemical systems or semiconductor devices. The voltages of saturated Weston cells unfortunately return to equilibrium very slowly after a thermal disturbance, so that great care is needed when they are transported between laboratories. Because of the large temperature coefficients of voltage characteristics of these cells, and especially because of the very large differential temperature coefficient between the two legs of a cell, very careful temperature control is required.

Recently it has become common to transfer the unit of voltage between laboratories by means of groups of standard cells contained in temperature-regulated enclosures which are maintained at constant temperature during shipment. A significant reduction in the uncertainty assigned to a transferred unit of voltage can be made by means of these systems. Unfortunately, with present enclosures the cell temperatures are not maintained as precisely as could be desired.

A great surge of activity is now taking place in many laboratories to determine whether or not the voltage steps induced by microwave radiation in the d-c current of a Josephson junction can be used to maintain the unit of voltage. Present indications are that Josephson junctions may well prove to be superior to saturated Weston cells for this purpose. An added attraction for this technique is that the use of Josephson junction voltage references would eliminate the need for transporting the unit of voltage between laboratories, since in principle only a published value for e/h and a reference to a time standard are necessary to determine the voltage at a given step of a Josephson junction. The major experimental problem encountered with Josephson junctions is that of accurately comparing their rather small voltages (presently of the order of 10^{-2} V) with the voltages of standard cells. The ability to make this comparison with high accuracy is essential in order to tie the voltage unit maintained with Josephson junctions to the absolute voltage unit obtained by means of a current balance.[2]

[2] The alternative possibility of *defining* the value of e/h would not eliminate the necessity for making measurements with current balances or electrometers, and would in fact degrade the accuracy of the ohm [15].

5. References

[1] Rayner, G. H., An absolute Determination of Resistance by Campbell's Metbod, Metrologia **3**, No. 1, 12–18 (Jan. 1967).

[2] Linckh, H. E., und Brasock, F., Eine Methode Zur Bestimmung des Widerstandswertes aus der Induktivität, Metrologia **4**, No. 3, 94–101 (July 1968).

[3] Page, C. H., A New Type of Computable Inductor, J. Res. Nat. Bur. Stand. (U.S.), **67B**, No. 1, 31–39 (Jan–Mar 1963).

[4] Homan, D. N., Some Techniques for Measuring Small Inductances, J. Res. Nat. Bur. Stand. (U.S.), **70C**, No. 4, 221–226 (Oct.–Dec. 1966).

[5] Thompson, A. M., and Lampard, D. G., A New Theorem in Electrostatics and its Application to Calculable Standards of Capacitance, Nature **177**, 888 (1956).

[6] Cutkosky, R. D., Evaluation of tbe NBS Unit of Resistance Based on a Computable Capacitor, J. Res. Nat. Bur. Stand. (U.S.), **65A**, No. 3, 147–158 (May–June 1961).

[7] Thompson, A. M., An Absolute Determination of Resistance Based on a Calculable Standard of Capacitance, Metrologia **4**, No. 1, 1–7 (Jan. 1968).

[8] Clothier, W. K., A Calculable Standard of Capacitance, Metrologia **1**, No. 2, 36–55 (April 1965).

[9] Vigoureux, P., A Determination of the Ampere, Metrologia **1**, No. 1, 3–7 (Jan. 1965).

[10] Driscoll, R. L., and Cutkosky, R. D., Measurement of Current with the National Bureau of Standards Current Balance, J. Res. Nat. Bur. Stand. (U.S.), **60**, No. 4, 297–305 (April 1958).

[11] Driscoll, R. L., Measurement of Current with a Pellat-Type Electrodynamometer, J. Res. Nat. Bur. Stand. (U.S.), **60**, No. 4, 287–296 (April 1958).

[12] Driscoll, R. L., and Olsen, P. T., Unpublished Communication to the Comite Consultatif D'Électricité.

[13] Huggins, R. W., and Sanders, J. H., Nuclear Magnetic Moment Ratios Measured in High and Low Fields, Proc. Phys. Soc. **86**, 53–63 (1965).

[14] Clothier, W. K., A Proposal for an Absolute Liquid Electrometer, Metrologia **1**, No. 4, 181–184 (Oct. 1965).

[15] Page, C. H., Letters to the Editor, Metrologia **1**, No. 2, 73–74 (April 1965).

DISCUSSION

U. Stille: May I just mention one point which touches the general title of our conference, fundamental constants. You mentioned regarding the realization of the ohm differing methods using inductance and using capacitance. There is in effect a difference between these two methods. If you are thinking about accuracies higher than 10^{-8} or in the region of 10^{-8}, then there is a special need. You see, inductance in our system of describing physics to which the SI units are appropriate enters only the magnetic constant which by the definition of the ampere indirectly got a fixed value. But regarding capacitance there enters the electric constant or, in other words, the square of the velocity of light. So then you must know the value of the velocity of light to this accuracy.

R. D. Cutkosky: That is correct.

Comment from the floor: Measure it with frequency. (*Laughter*)

A Stark Voltmeter as an Electromotive Force Reference

Ko Hara, Takeshi Kobayashi, Toshimi Matsui, Takahiro Nakase, and Genta Yonezaki

Electrotechnical Laboratory 5-4-1, Mukodai, Tanashi, Tokyo, Japan

An experimental study of an electromotive force reference based on the Stark effect in a molecular rotational transition has been made. The spectrometer used was a Stark type with a Fabry-Perot resonator and provided a uniform field by means of a transparent electrode in the resonator. The example chosen was the $J=2\leftarrow1$, $K=1\leftarrow1$, $\Delta M=\pm1$ transition of the symmetric rotor *s*-trioxane ($C_3H_6O_3$). The rotational constant B_0 (uncorrected for centrifugal distortion) and the dipole moment were found to be 5273.2503 ±0.0006 MHz and 2.06430 ±0.00016 debye respectively. The reproducibility of the system as a voltmeter is estimated to be 80 ppm at 350 V/cm.

Key words: Fabry-Perot resonator; microwave spectrometer; molecular constants; Stark voltmeter.

1. Introduction

The present work is an experimental study of the possible use of the Stark effect of a molecular rotational transition as an electromotive force reference. A Stark cell consisting of a modified Fabry-Perot (FP) reflection resonator with a transparent Stark electrode between the reflectors has been built, and its performance has been surveyed using the transition ($J=2\leftarrow1$, $K=1\leftarrow1$, $\Delta M=\pm1$) of *s*-trioxane, $C_3H_6O_3$. A preliminary value of the dipole moment of this molecule has been determined as 2.06430± 0.00016 debye in a field of about 350 V/cm.

2. Stark Splitting and Sample Gas

Stark splitting formulae found in texts are generally second order approximations and are not adequate for the present purpose. New formulae have been worked out [1] following Shirley [2]. The results for the transition $J=2\leftarrow1$, $K=1\leftarrow1$, $\Delta M=\pm1$ of a symmetric top molecule are as follows.

$$S_0: \quad f_0(\lambda)=4B \tag{1}$$

$$S_1, S_2: \quad f_1(\lambda)=f_2(-\lambda)=B(4+0.1666667\lambda +0.0163360\lambda^2-0.0030883\lambda^3 +0.0001855\lambda^4\pm0.0001177\lambda^4) \tag{2}$$

$$S_3, S_4: \quad f_3(\lambda)=f_4(-\lambda)=B(4+0.1666667\lambda +0.0536376\lambda^2+0.0026547\lambda^3 -0.000149\lambda^4\pm0.000189\lambda^4) \tag{3}$$

$$S_5, S_6: \quad f_5(\lambda)=f_6(-\lambda)=B(4+0.5000000\lambda +0.0494048\lambda^2-0.0031250\lambda^3 +0.000049\lambda^4\pm0.00015\lambda^4). \tag{4}$$

Here S_0 represents the zero-field transition frequency; S_1, S_2 refer to the $M=\pm1\leftarrow0$ Stark components; S_3, S_4 to the $M=\pm2\leftarrow\pm1$ components; and S_5, S_6 to the $M=0\leftarrow\pm1$ components.

The splitting pattern is given in figure 1 together with the relative intensities. In the above equations, $\lambda=\mu E/hB$, where μ is the dipole moment, E is the applied d-c field, h is Planck's constant and B is the rotational constant. When *s*-trioxane is used as a sample gas, $\lambda=1$ corresponds to about 5,000 V/cm. Lines S_1 and S_2 have been studied. S_3 and S_4 are so close to S_1 and S_2 (respectively) at low field that they

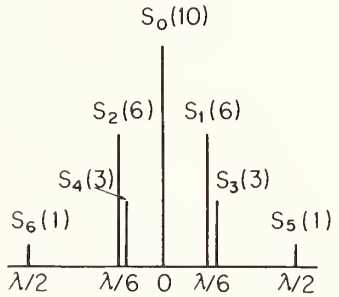


FIGURE 1. *Stark splitting pattern of symmetric top molecule* ($J=2\leftarrow1$, $\Delta M=\pm1$).

cause some difficulty in determining the center frequencies of S_1 and S_2.

S-trioxane has been selected for study because of its relatively large dipole moment, absence of quadrupole hyperfine structure, and practically pure natural abundance of each constituent atom, and also because it has a rotational frequency within the range of the microwave components which we had at the beginning of this study. Several molecules other than *s*-trioxane having rotational frequencies in the 20 GHz range are also adequate for the

Table 1. *Spectroscopic data of some symmetric top molecules*

Molecule	Rotational constant B (MHz)	Dipole moment μ (debye)
PF_3O	4594.25	1.77
CH_3CF_3	5185	2.32
$C_5H_6O_3$	5273.6	2.08

purpose, but are difficult to purchase. Some parameters of these molecules are given in tables 1 and 2.

The absorption coefficient γ_{max} without dc field has been estimated to be about 1.8×10^{-6} cm^{-1}, and the line width Δf has been estimated to be $25p$ MHz (p is the pressure in torr). γ_{max} begins to decrease at pressures below 0.01 torr; accordingly, the optimum sensitivity would be attained at about 0.005 torr. A receiver of noise figure 15 dB would resolve 1 percent to 0.1 percent of the line width, and thus an applied dc field of 1000 V/cm would be resolved to about 1 ppm.

3. Apparatus

3.1. Stark Cell

Instead of a traditional waveguide Stark cell, a modified Fabry-Perot (FP) reflection type resonator cell with a transparent Stark electrode between the reflector pair has been used. A FP resonator cell has previously been tried by Beers and Russel, [3], but a transparent electrode was not used. This electrode enables us to establish an extremely uniform d-c field and to adjust the electrode separation independently of the resonant condition to some extent.

The preliminary Stark cell is shown in figures 2 and 3. Plate A is a fused quartz optical flat of 300 mm diameter, 30 mm thickness, and one half fringe flatness, on which a silver film about 3,000 Å thick is evaporated. This is used as a perfect reflector of the FP resonator and a dc electrode simultaneously. Plate B is similar to A but has an evaporated tellurium film about 200 Å thick. The surface resistance of this film is about 100 kΩ and it is almost transparent at 20 GHz. Plate C is made of silvered stainless steel and its diameter is 180 mm. At the center of this plate there is a 7 mm diameter aperture to couple microwave power from the waveguide behind it. This aperture is replaceable so

Table 2. *Data required for hfs of materials in table 1*

Atom	Z	A	Natural abundance (%)	Spin	Quadrupole moment
H	1	1	99.9851	1/2	None
C	6	12	98.892	0	None
O	8	16	99.758	0	None
F	9	19	100	1/2	None
P	15	31	100	1/2	None

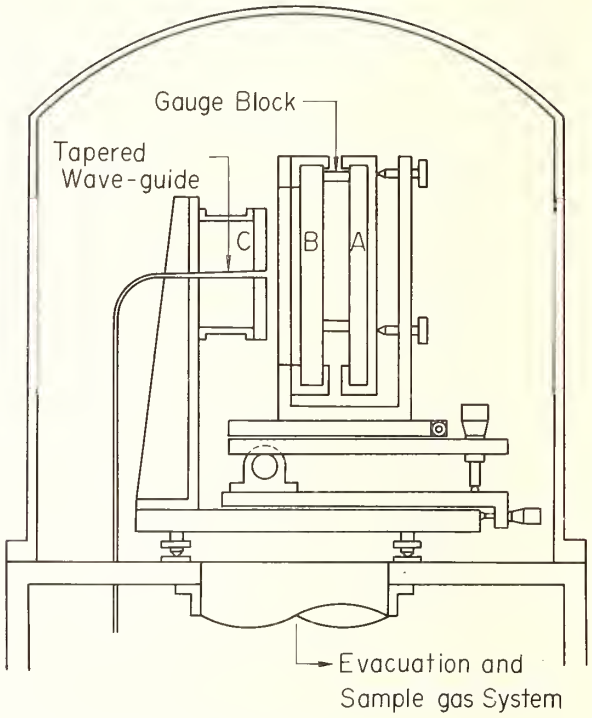


Figure 2. *Structure of the preliminary Stark cell.*

that the FP resonator coupling can be adjusted near critical. The separation and parallelism of A and B are fixed by three identical 28.4447 mm gauge blocks. The resonant performance of the structure as a spectrometer has been analyzed fully by one of the authors [4]. The whole system is mounted on a stainless steel base plate and covered by a 70 cm diameter stainless steel bell jar.

3.2. Microwave Circuit

Figure 4 shows the microwave circuit of the spectrometer with Stark modulation. A 400 MHz signal generator is used as a master oscillator. The output frequency of the 400 MHz signal generator is

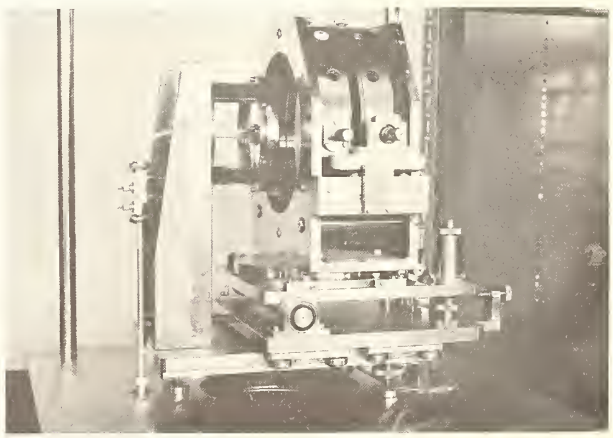


Figure 3. *Photograph of the Stark cell with bell jar removed.*

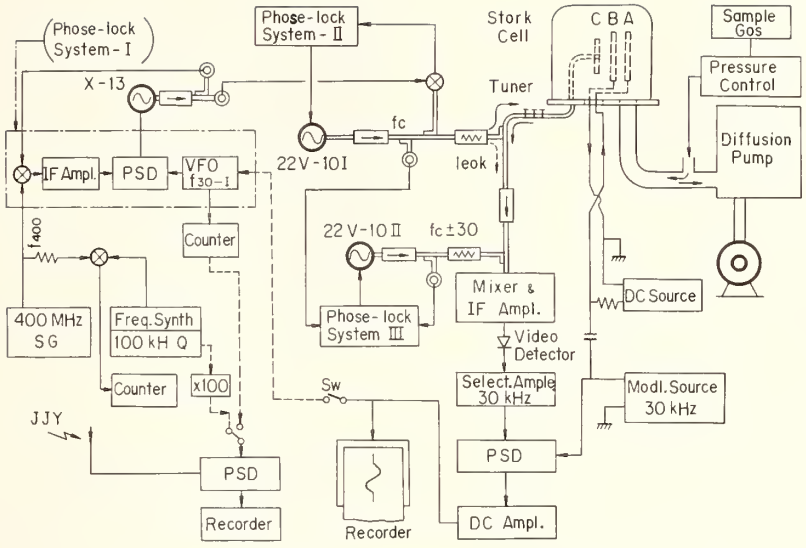


FIGURE 4. *Block diagram of the spectrometer.*

monitored by harmonic mixing with a frequency synthesizer and is adjusted manually to keep the beat frequency within 10 Hz. An X-13 klystron is phase-locked by phase-lock system I to the 400 MHz signal generator and yields a 10 GHz output. A 22V-10 klystron (I) is then phase-locked to the X-13 and yields a 20 GHz output which is coupled into the cell. Another 22V-10 klystron (II) is phase-locked to the first 22V-10 and provides the local oscillator output for a superheterodyne receiver.

The spectra are observed by applying to the Stark electrode a dc voltage on which a 30 kHz square wave modulation and a sweep voltage are superposed. The reflected wave from the cell is then mixed with the local oscillator signal and amplified by an IF amplifier. The signal then passes through a video detector where the modulation signal is recovered and detected by a phase sensitive detector. The output of the phase sensitive detector is amplified by a dc amplifier and recorded on a strip chart recorder.

To measure the center frequency of the spectrum, a signal from the output dc amplifier is fed back to a voltage tunable oscillator (VFO) which serves as the IF reference of phase-lock system I. In this AFC mode of operation, the output frequency of 22V-10 I is locked to the zero derivative point of the absorption line and is determined in terms of the synthesizer setting, the IF reference frequencies of the phase-lock systems and the beat frequency between the 400 MHz signal generator and the synthesizer. Any oscillator working in this system can be calibrated against JJY (the standard frequency broadcast station in Japan) with an accuracy on the order of 1 part in 10^8.

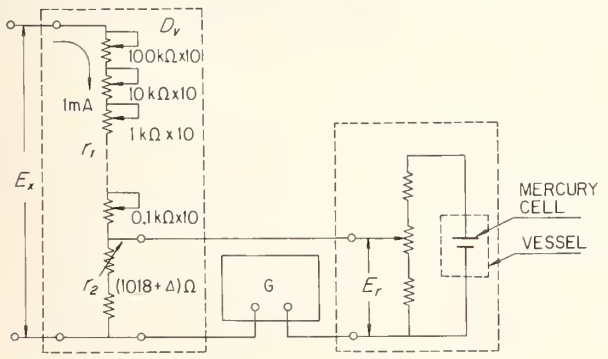


FIGURE 5. *Voltage calibration circuit.*

3.3. Voltage Calibration

A seven decade voltage reference with range to 1000 V is used as a voltage standard. It is stable within ±1 ppm for about 20 min and is satisfactory for the present purpose. Its output voltage E_x is calibrated against a standard cell using the resistive divider D_v shown in figure 5. E_x is given by

$$E_x = (r_1 + r_2)/E_r/r_2.$$

4. Stark Coefficients and Center Frequency Measurement

4.1. Stark Coefficients

Equation (2) (for example) can be rewritten as a power series in the applied voltage, and the coefficient of each term (Stark coefficient) expressed as follows:

$$a_1 = (1/6)\,(\mu/hD) \tag{5}$$

$$a_2 = (8{,}892/15{,}120)\,(a_1^2/B) \tag{6}$$

$$a_3 = (363{,}096/544{,}320)\,(a_1^3/B^2) \tag{7}$$

where D is the separation between the electrodes. B and a_1 are found to be

$$a_1 = (f_1 - f_2)/2V_s - a_3 V_s^2 \tag{8}$$

$$B = (f_1 + f_2)/8 - a_2 V_s^2/4 \tag{9}$$

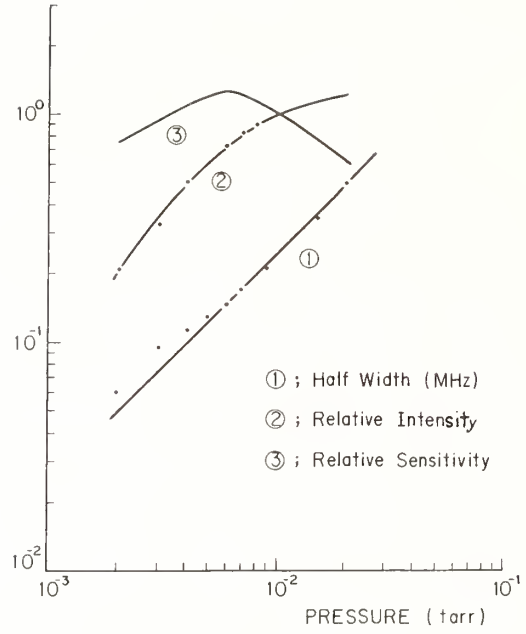


FIGURE 6. *Relative sensitivity of s-trioxane versus pressure.*

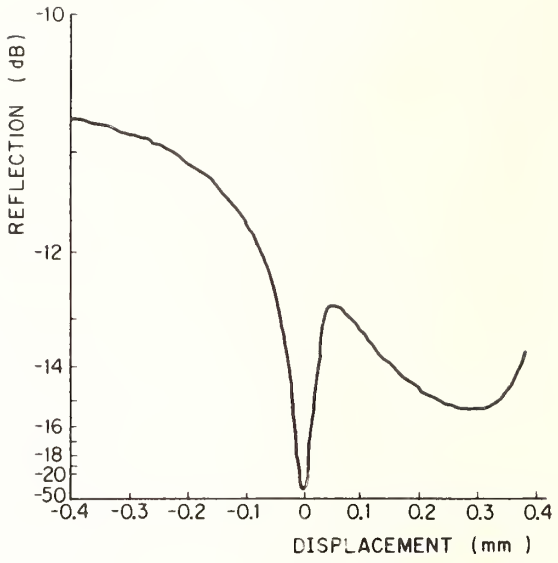


FIGURE 8. *Reflection characteristic of the FP resonator.*

where f_1 and f_2 are the measured center frequencies of S_1 and S_2 respectively, and V_s is the applied voltage. B and a_1 can be obtained by successive approximation, because the second terms of eqs (8) and (9) are small correction terms.

The zero derivative frequencies we measure differ from the center frequencies for the following reasons:

(1) Detuning of the cell.
(2) Leakage coupling between the microwave source and the mixer at the directional coupler.
(3) Satellite lines S_3 and S_4, which are close to S_1 and S_2 respectively. Brief explanations will be given for these in the following.

4.2. Factors Affecting the Center Frequency Measurement

As one can see in figure 6, optimum resolution is reached at about 0.006 torr, where the line width is about 150 kHz. The Q of the FP resonator is measured to be about 1,500. Accordingly, the line width is not narrow enough to neglect the so-called Q-pulling [4]. Numerically, a deviation of the FP reflector separation of 0.1 μm from the resonant condition causes a deviation of 2.5 percent of the line width, i.e., 4 kHz. The correct resonance condition can be established by observing insensitivity of the reflected power from the resonator to forced mechanical vibration of the cell.

Leakage power at the directional coupler in front of the cell from 22V-10 I to the receiver results in an error in the measured frequency, unless the phase of the reflected wave is exactly in phase or out of phase with the leakage wave. This error can be avoided by balancing out the leakage power as far as possible and adjusting the reflected power to be 35 dB or more above the leakage power. This adjustment causes little loss in sensitivity due to mismatch of the resonator.

A satellite line, e.g. S_3 for S_1, has a somewhat complex influence on the center frequency measurement of the main line. Its effect depends on the matching condition of the resonator, applied dc field, amplitude of Stark modulation signal and line width. When the reflected amplitude from the cell is ad-

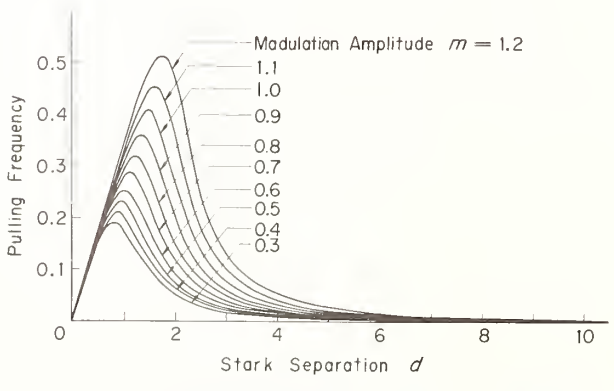


FIGURE 7. *Satellite line pulling as a function of electrode separation and modulation.*

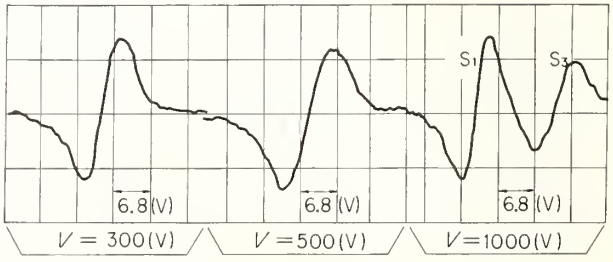


FIGURE 9. *Profiles of the spectra.*

justed to be more than 100 times the absorbed amplitude in the cell, the influence of the resonator matching condition can be neglected. The pulling by the satellite line is calculated for this case and the result is shown in figure 7. In this figure, the Stark separation d between the main and satellite line, the modulation amplitude m (converted into frequency), and the pulling frequency are all normalized to the line width. Accordingly, a pressure measurement of the sample gas is essential to obtain the center frequencies of S_1 and S_2. The necessary correction is obtained by a numerical calculation similar to that on which figure 7 is based.

5. Experimental Results

The operational conditions for the spectrometer are as follows: The noise figure F of the receiver is about 15 dB with 1 mW local oscillator power. The input signal power to the cell (about 300 μW) is stable within 0.1 percent. The separation between plates C and B is 5 mm and corresponds to the lowest order resonance of the resonator. The reflection characteristic versus displacement of A-B against C is given in figure 8. It shows that higher modes are not well separated and that the loss at the dominant mode is about 2 percent/transit.

Dispersion type spectra for S_1 thus obtained are shown in figure 9 at about 300, 500 and 1000 V respectively. S_1 and S_3 begin to separate at 500 V and the profile becomes broad and asymmetric; they do not separate appreciably at 300 V. At 1000 V, they separate clearly and the line width is restored, but the profile remains slightly asymmetric.

An example of the scatter of a measured zero derivative frequency (AFC mode of operation with an averaging time of 10 s) is shown in figure 10 at about 300, 500 and 1000 V respectively. The standard deviation for each run is about 1.5 kHz; this corresponds to about 8 mV/cm.

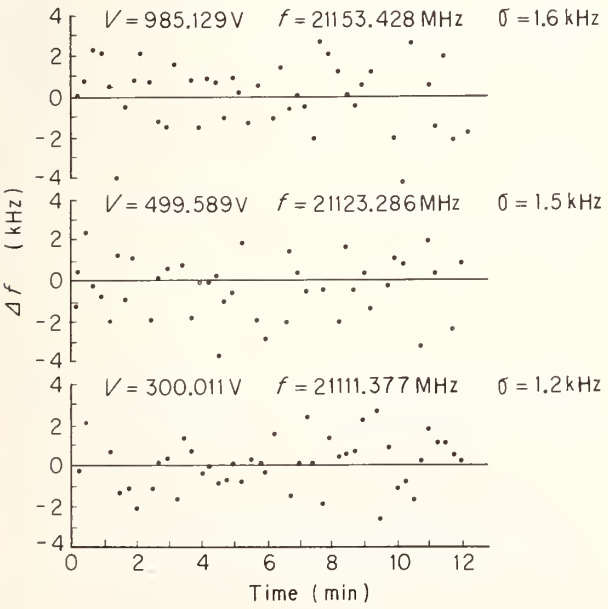


FIGURE 10. *Resolution and short term stability of the spectrometer.*

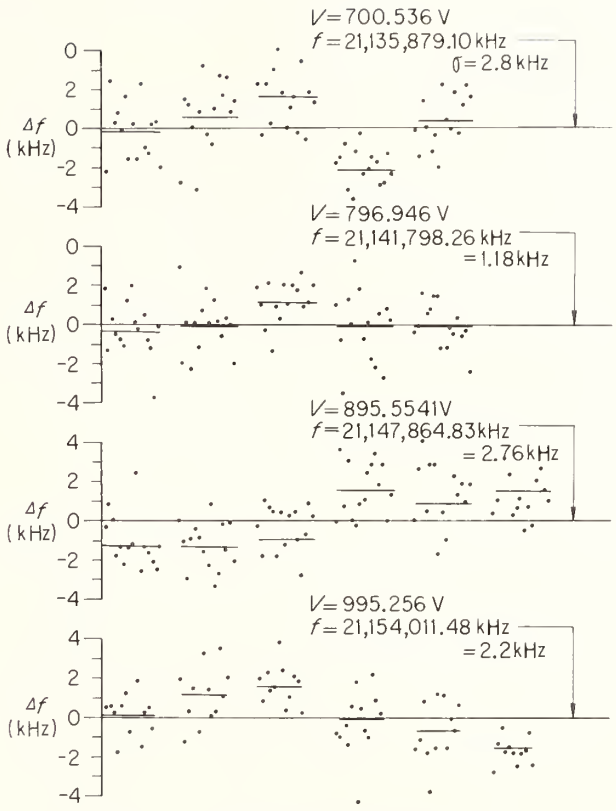


FIGURE 11. *Effect of resetting of the resonator on measured frequency.*

Figure 11 shows the effect of repeated setting of the reflector separation on the measured frequency for S_1. Each horizontal solid line groups together measured values with a common reflector setting and represents their average. The standard deviation for each applied voltage is about 2.5 kHz, about 1.7 times that in figure 10. This fact shows that the error due to resetting, i.e. Q-pulling error, exceeds that due to receiver resolution.

Following the procedure in section 4.1, B and μ/h are determined at about 700, 800, 900, and 1000 V using S_1 and S_2. The corrections due to pulling by satellite lines are 19, 8, 4, and 2 kHz for 700, 800, 900, and 1000 V respectively. The results are given in table 3, where the weight for each applied voltage is taken to be inversely proportional to the estimated variance of B and μ/h. The variance of B is greater for lower applied voltage because the estimated error for satellite line pulling is greater. The variance of μ/h is greater for lower applied voltage because the Stark splitting for lower voltage is smaller. The final estimates of B and μ/h are weighted means obtained using the weights given in the table. The quoted standard deviations are measures of their scatter for different applied voltages.

TABLE 3. *Rotational constant and Stark effect constant of* s-*trioxane* (S_1 *and* S_2)

Voltage (V)	Rotational constant B (kHz)		Stark effect constant μ/h (esu)	
	$B\pm\sigma$	Weight	$\mu/h\pm\sigma$	Weight
			$\times 10^8$	
1,000	5,273,250.0±0,5	1	3.11549±0.00015	1
900	5,273,251.0±0.8	0.4	3.11519±0.00018	0.7
800	5,273,250±1	0.25	3.11550±0.00015	1
700	5,273,252±3	0.03	3.11490±0.00025	0.36
Weighted average	5,273,250.3±0.6		3.11536±0.00024	

6. Discussion and Conclusion

The final quoted standard deviation of B suggests that the center frequency measurement has been done with a standard deviation of 2.4 kHz. This is compatible with that due to resetting error and corresponds to a 12 mV/cm electric field resolution. This is about 35 ppm for an applied field of 350 V/cm and implies that the standard deviation of μ/h may be 0.00011×10^8 (esu). The estimated standard deviation (table 3) is about twice this value. This discrepancy suggests that there is a large non-reproducibility in the electric field of the order of 70 ppm. The existence of such a large variation is almost inconceivable at first sight. However, it may be explained as follows: Measurements for each value of applied voltage were done in the course of a single day. Measurements for different voltages were separated by intervals of several days. The experimental conditions were found to be fairly stable during the course of a single day. During the period between successive measurements, some change may have occurred in the electrode separation or in the effective voltage between the electrodes. Thermal instability of the electrode separation may be possible because of the complexity of the cell construction and interruption of the laboratory air conditioning during the night. Alternatively, a difference between the effective electrode voltage and the electrode terminal voltage may be the origin, because the surface resistance of the transparent electrode is relatively high. These problems should be attacked in the next step of the study.

The ultimate accuracy which will be reached (limited only by the detector resolution in the preliminary spectrometer with s-trioxane) is thus estimated to be about 1.5 kHz. This is equivalent to 7.5 mV/cm and is about 20 ppm at 1000 V, as compared with the present accuracy of about 77 ppm.

The difference of the work functions between the two electrode materials has been measured to be 24±10 mV and is negligible at this stage of the study.

Future improvements will include (1) use of methyl cyanide (CH_3CN) as a sample gas and (2) use of a Schottky barrier diode as a detector. According to Beers [5] the electric field resolution $\delta E/E$ is

$$\delta E/E \propto (\gamma E)^{-1}(FB\nu/QV)^{1/2} \qquad (10)$$

where the notation is self-evident except that V stands for the effective volume of gas. Assuming $\gamma=1.2\times10^{-4}$ cm^{-1} for CH_3CN, as compared with $\gamma=2\times10^{-6}$ cm^{-1} for $C_3H_6O_3$, the final resolution which can be reached with CH_3CN may be better than 0.5 ppm. Furthermore, another factor of 2 or so may be obtained by reducing the receiver noise figure.

Table 4 summarizes the constants for s-trioxane which have appeared in the literature together with the results of the present work.

TABLE 4. *Summary of* B *and* μ *of* s-*trioxane*

B (MHz)	μ (debye)	Worker
5,273.6	2.08±0.01	E. Amble, [6]
5,273.23±0.01	2.07±0.04	T. Oka et al., [7]
5,273.25		J. Bellet et al., [8]
5,273.2503±0.0006	2.06430±0.00016*	Present work.

* Planck's constant $h=6.626196\times10^{-27}$ erg·s is used.

7. References

[1] Hara, K., Denkishikenjo Iho **31**, 1136 (1967) (Bulletin of the Electrotechnical Laborarory, **31**, 1136 (1967)).
[2] Shirley, J. H., unpublished report (Dec. 9, 1960).
[3] Beers, Y., and Russel, T. W., IEEE Trans. on Instr. and Meas. **IM-15**, 380 (1966).
[4] Kobayashi, T., Denkishikenjo Iho **32**, 239 (1968) (Bulletin of Electrotechnical Laboratory, **32**, 239 1968).
[5] Beers, Y., Rev. Sci. Instr. **30**, 9 (1959).
[6] Amble, E., Phys. Rev. **83**, 210A (1951).
[7] Oka, T., et al, Bull. Chem. Soc. Japan **37**, 4 (1964).
[8] Bellet, J., et al, C. R. Acad. Sc. (Paris) **262**, Serie B, 885 (1966).

DISCUSSION

V. W. Cohen: What was the nature of your voltage reference throughout this work?

K. Hara: As the high voltage source, a commercial seven-position voltage source is used, and that is calibrated in terms of standard cells by means of a resistive divider.

U. Stille: I have seen your abstract. You have given some experimental values, for instance, for the Stark effect constant. You have given it in cgs units but with an accuracy of one or two parts in 10^5. Which value of the elementary charge have you used?

K. Hara: To convert μ/h from SI units to cgs esu, the elementary charge does not enter. In this experiment, the perturbation energy which is approximately proportional to μE is measured in terms of frequency ν. Thus, μ/h is determined in units of $(\mathrm{Hz \cdot m})/\mathrm{V}$, when SI units are used for the electric field E and the frequency ν. To convert this value into cgs esu, only the velocity of light, c, is necessary except a numerical factor. The value of h is necessary, to calculate the dipole moment μ. Planck's constant is found in the 1969 adjustment table of the fundamental constants. Although μ is defined as the product of charge and length, the experimental situation is like this.

An All-Purpose 8-Place Potentiometer for Ultra-Precise Measurement

Leon W. Dean

Dean Laboratory, Mansfield, Massachusetts 02050

This paper describes a wide range double potentiometer capable of measuring two emf's in rapid succession with a dial resolution of 1 part in 10^8 and an accuracy limited by the accuracy with which the ratio of standard cell resistor value to range resistor value can be determined. Other limits are stability of current source and ultimately, as in any measuring circuit, circuit noise and thermal emf's. The basic circuit is a Lindeck element in which the current supplied to the range resistor is obtained by dividing the output of a constant current source by means of a precision, double, eight-dial Kelvin-Varley divider. This divider-Lindeck element combination is sometimes referred to as a Julie element. The instrument has built-in facilities for self-checking and for adjusting critical resistance ratios to achieve divider linearity. (For detailed descriptions of this instrument, see [1, 2]).

Key words: DC measurements; eight place double potentiometer.

References

[1] Julie, L., Electronic Instrument Digest **6**, #10, 47 (October 1970).

[2] Dean, L. W., and Ludwig, D. R., Measurements and Data **4**, #1, 74 (Jan.–Feb. 1970).

PROTON GYROMAGNETIC RATIO

The Gyromagnetic Ratio of the Proton: A Survey

P. Vigoureux

National Physical Laboratory, Teddington, Middlesex, England

The gyromagnetic ratio of the proton can be determined in terms of the ampere by two methods: one uses a weak magnetic field, the other a strong magnetic field. Either method used singly yields the constant if a current balance or other dynamometer is also available, but if equipment for both methods is provided, the dynamometer can be dispensed with, and not only the gyromagnetic ratio, but also the ampere, obtained.

There may be advantages therefore in building equipment for both methods, especially in laboratories which do not possess a current balance. As however one of the operations in the strong-field method consists in "weighing" a current, the final result is liable to approximately the same uncertainty which affects the weak-field method combined with a current balance.

Key words: Absolute determination of ampere; gyromagnetic ratio of the proton; high field method; low field method.

Physicists want to know the gyromagnetic ratio of the proton as accurately as possible because it influences evaluation of a number of other important constants of physics. To electrical engineers and others interested in realizing and preserving electric units it is an important quantity because the apparatus needed to determine it is also suitable for maintenance of the ampere. In this survey I propose to define the constant, give general information on it, describe the methods of measuring it, compare their

advantages, and discuss their application to the conservation and possibly to the realization of the ampere.

We call the gyromagnetic ratio of a particle the quotient m/L of its magnetic moment m and its angular momentum L, and we denote it by γ. According to the principles of quantum mechanics the axis of spin of a particle in a magnetic field of flux density B is aligned either parallel or antiparallel to the direction of B. If the magnetic moment is m the magnetic energies in the two positions are $-mB$ and $+mB$, and a transition from one state to the other corresponds to a change of energy $2mB$. Each transition is accompanied by emission or absorption of radiation of energy equal to the difference of the energies of the two states. If the frequency and angular frequency of the radiation are denoted by f and ω, and Planck's constant by h, the energy absorbed or emitted is also equal to hf or $h\omega/2\pi$. Equating this expression to $2mB$, and remembering that the angular momentum L of the particle is $h/4\pi$, we find

$$\omega = \gamma B. \qquad (1)$$

The classical picture, easier to visualize and equally instructive, is that of a particle possessing mass and charge, and spinning about an axis. The mass and spin endow it with angular momentum L and the charge and spin with magnetic moment m. The axis of the spinning particle, inclined at some angle θ with the direction of the magnetic field B, experiences a torque $mB\sin\theta$, but the angular momentum prevents it from at once aligning itself with the field, and causes it to precess about the field, as a spinning top does about the vertical. It can be shown that the angular frequency of this "Larmor precession" is given by eq (1) above.

The ratio m/L would be equal to $e/2M$ if to each element of mass there were a proportional element of charge, but as for the proton the measured value of γ is about 5.58 times that predicted, we are forced to abandon our over-simple picture of the particle and regard γ as an experimental constant to be measured for each nucleus.

If the protons are provided by matter in bulk, if for example they are nuclei of hydrogen atoms in the molecules of a volume of water, collisions are caused by the thermal energy, which is $kT/2$ for each degree of freedom and far exceeds the magnetic energy mB even in strong fields. The collisions produce a decrease, on the average, of the angles of precession, and force the axes of spin to set themselves parallel to the field. The process is subject to a relaxation time which differs from liquid to liquid, and is about 3 seconds for water.

Formula (1) suggests that to measure the gyromagnetic ratio we must place some substance containing protons, preferably one rich in protons, like water, in a known magnetic field, somehow arrange for precession to occur, and observe the frequency of the emf which the precessing protons induce in a coil surrounding the substance. The classical picture

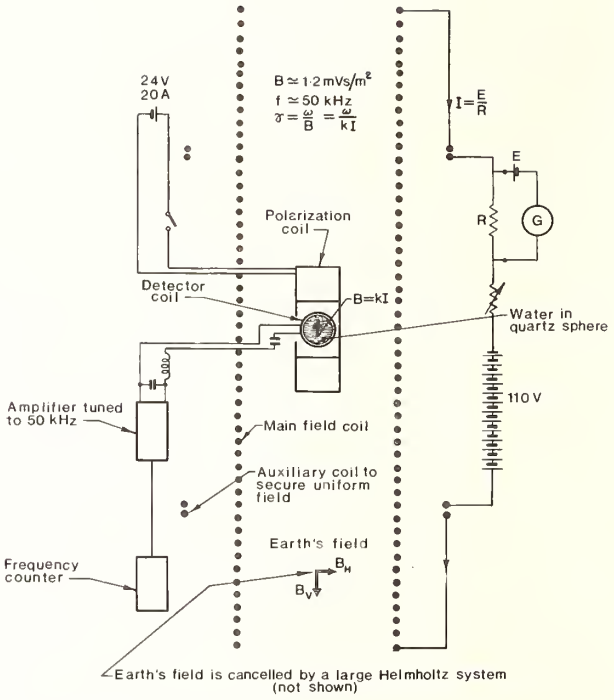


FIGURE 1. *Measurement of gyromagnetic ratio of the proton by the method of free precession in a weak magnetic field.*

of the phenomenon moreover tells us what the orientation of the detector coil should be.

In figure 1 the magnetic field is produced by the long coil and by an auxilliary coil of a few turns which makes up for the finite length, usually about 1 m, of the long coil, and ensures uniformity of field throughout the volume of the water container, a spherical glass or quartz shell 40 mm in diameter at the center of the apparatus. Round the sphere is a detector coil tuned to the frequency appropriate to the magnetic field, about 50 kHz; the tuned circuit is followed by a narrow-band amplifier and a counter to measure the time of 10, 20, or 40 thousand periods of the signal. The more uniform the field, the longer is the decay time, and the larger therefore is the number of periods that can be measured, and the higher the accuracy of the measurement. The current in the field-forming coils flows also through a resistor R of nominal value 1 ohm and is adjusted automatically or by hand for the p.d. at the potential terminals to balance the emf of the Weston cell E. The current is thus nominally E/R but if a current balance or other dynamometer is available E/R can be determined in amperes.

A few seconds after the standard field has been established the protons will have become on the average aligned with it and there is no signal. The largest signal is obtained if to start with the axes of spin are at right angles to the standard field and if the polarization of the protons is large. The large polarization is produced by submitting the sphere to a high magnetic field for a few seconds. This opera-

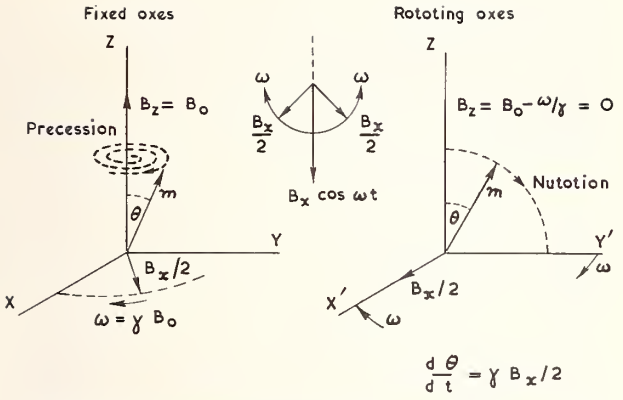


FIGURE 2. *Nutation caused by field rotating at precession frequency.*

tion can be performed by placing the sphere in a permanent magnet a distance away, or by means of an extra coil at the center of the apparatus. Whichever method is used, the time taken to bring the sphere to the center of the apparatus or to switch off the polarizing current allows the polarization to rotate and assume the direction of the standard field. We then rotate it one right angle by sending a pulse of current at precession frequency either in the detector coil or in an auxilliary coil. This pulse is produced by a generator or simply by connecting a charged capacitor to the coil; it causes nutation and is adjusted to last long enough for the nutation angle, θ in figure 2, to reach a right angle. As removal of the polarizing field and application of the nutation pulse occupy a comparatively short time, the polarization is still high. The precession which then takes place gives rise to a large signal, and although it dies down in one or two seconds, there is plenty of time to measure the period.

The standard field B is calculated from the current and the dimensions of the two field-forming coils, which must be wound in single layers if their dimensions are to be measured accurately. The Earth's field must be removed or allowed for. A simple way of eliminating it is to maintain the requisite current in two pairs of Helmholtz coils, one with axis vertical, the other with axis along the horizontal component of the Earth's field. Cancellation need not be perfect for if there are residual components B_1 along and B_2 perpendicular to the standard field, B_1 is in any case eliminated by reversing B and taking the mean of the two frequencies, whereas B_2 gives rise to a resultant field differing in magnitude as the hypotenuse differs from the long side of a right-angled triangle of small third side, i.e., by a relative amount $B_2^2/2B^2$. This ratio is less than 1 in 10^7 when B corresponds to about 50 kHz, even if the cancellation by the Helmholtz coils is good to only 1 percent of the Earth's field. In practice cancellation to 1 in 1000 is easy.

The value γ obtained in this way is not the gyromagnetic ratio proper, which refers to an isolated proton, because the electron of the hydrogen atom screens the proton from the field B. The difference of about 26 in 10^6 has been calculated and also verified by experiment. For maintenance of the ampere it is neither necessary nor even desirable to apply this correction, nor is a correction needed for the paramagnetism of the liquid and the shape of the container—it is simpler to agree to use water in a spherical vessel, so that the results of different laboratories may be strictly comparable.

Measurement of γ in terms of E/R by this "weak-field" method is subject to a probable uncertainty of perhaps 1 in 10^6, mostly due to systematic errors which do not come in when the apparatus is used to test the constancy of E/R. For this purpose it is not even necessary to have previously determined γ to the above accuracy, it is sufficient to agree on a value, or just to state the particular value used in any one measurement. The value appears to be $2.675\,127 \times 10^8$ rad/sT, the units being derived from eq (1); an equivalent form sometimes useful is 42.576 MHz/T, the units in this case being derived from the actual method of measurement. Whichever form is used, it is important to remember that γ depends on the unit of current, not on the unit of electric potential.

A determination of γ in terms of the ampere, rather than in terms of the units of emf and resistance of the laboratory as maintained by Weston cells and standard resistors, is less accurate, for the ampere itself is subject to a probable error of some 4 parts in 10^6. It may be asked why that should be so, since in each case similar coils of wire wound on cylindrical formers are employed. The answer is that calculation of magnetic flux density at the center of a long coil is inherently more accurate than that of force between two coils, because the latter is very dependent on the diameters as well as the pitch of the windings, whereas the former involves mainly the pitch, with the diameter appearing only in so far as the coil is not infinitely long. The outcome is that the ampere can be preserved more accurately than it can be realized.

The gyromagnetic ratio of the proton can also be determined by a method known as "strong-field," which makes use of apparatus normally employed to study nuclear magnetic resonance. The magnetic field of about 1 T is provided by the usual large magnet, more often an electromagnet. The water container, made much smaller, say 10 mm instead of 40 mm in diameter, is placed in the center of the gap as in figure 3. Rather than observe free precession it is more convenient in this method to stimulate transitions by a continuous-wave radiofrequency field, of approximately 42.5 MHz if the flux density is 1 T. The exciting coil surrounds the water container and its axis is parallel to the pole pieces. When the applied radiofrequency is equal to the precessional frequency the transitions are a maximum, in other words resonance occurs, which can be detected either by the reaction on the exciting circuit or else by observing the signal induced by the precessing

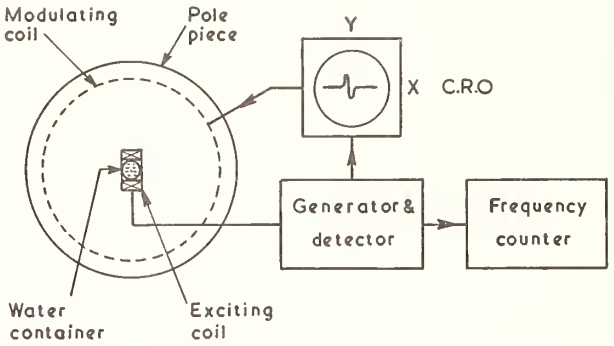


FIGURE 3. *Measurement of gyromagnetic ratio of the proton in a strong magnetic field.*

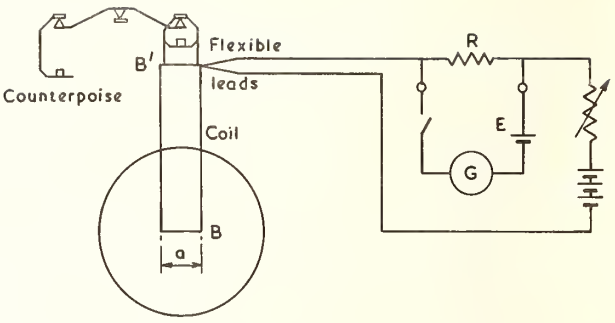


FIGURE 4. *Determination of flux density at center of gap of magnet used for measuring precessional frequency of protons.*

protons in a second coil also surrounding the container, but orthogonal with the exciting coil and the magnetic field B. Since the signal is derived from a continuous-wave field, measurement of its frequency presents no problem, but it is more convenient to keep the frequency fixed at approximately the correct value and to modulate the magnetic field of the magnet at a low frequency, 10–50 Hz, by sending current in two coils placed inside the gap against the pole faces. Alternatively modulation can be effected by a uni-directional current increasing linearly with time, as provided by the time base of oscilloscopes. The signal is displayed vertically on the screen of the oscilloscope and the modulation current horizontally. The small correction to B is estimated from the position of the resonance curve along the time base.

In the strong-field method the magnetic flux density B, which must be known in order that γ may be obtained from eq (1), cannot be calculated, it must be measured. The measurement is effected by observing the force which the field in the gap exerts on a known current. A coil of rectangular shape, long enough for one short side to be well above the pole pieces when the center of the opposite side lies in the region of the gap previously occupied by the water container, is suspended from one end of a balance beam whose other end carries a counterpoise. The change of force due to reversal of current in the coil is measured by means of a weighing mass removed from, or added to, a scale pan hanging from the same suspension as the coil. The current is controlled by a Weston cell of emf E and a resistor of resistance R, and is thus equal to E/R in terms of the laboratory units. The change of force on reversal of the current is $2aN(B-B')E/R$ where N is the number of turns of the coil, and B' is the small stray flux density at its upper end, figure 4. This stray field is measured by a magnetometer, the field in the gap is mapped by relative precessional frequency measurements, averages are taken, and the final result is referred to the flux density B at the lower side of the rectangle. As before, if a current balance or dynamometer is available, E/R can be determined in amperes, B deduced from it as above, and γ obtained from eq (1).

If however, instead of the weak-field or the strong-field apparatus and a dynamometer, the weak-field and the strong-field apparatus are both available for measurement of γ in terms of the same E and R, we are presented with an interesting possibility. The weak-field method gives

$$\begin{aligned}\omega_W &= \gamma B_W \\ &= \gamma \mu_0 H \\ &= \gamma \mu_0 k I \\ &= \gamma \mu_0 k E/R\alpha \end{aligned} \qquad (2)$$

where α is the ratio, close to unity, of the ampere and the laboratory unit maintained by means of resistors and Weston cells. The strong-field method gives

$$\begin{aligned}\omega_S &= \gamma B_S \\ &= \gamma k'/(E/R\alpha). \end{aligned} \qquad (3)$$

In these equations k depends on the dimensions of the field-forming coils, k' on the dimensions of the rectangular coil, the weighing mass, the acceleration due to gravity, and the correction for the stray field B'.

From eq (2) and eq (3) we obtain not only γ by multiplication, but also α by division, and we are able to dispense with the dynamometer, yet realize the ampere as well as determine the gyromagnetic ratio of the proton. The scheme is so attractive that it is worth considering what we gain and what we lose by adopting it. The design and development of a current balance is a major undertaking and even if good use is made of accumulated experience and of improvements in materials, particularly for coil formers, it is not easy to reduce the probable uncertainty of measurement of E/R by a balance to 2 in 10^6. The balance and the weak-field apparatus in conjunction would then yield γ to about 3 in 10^6.

The uncertainty of the strong-field method is partly due to uncertainty of the effective field B for use in eq (1). What is needed is B at the center of the container. It is not very difficult to arrange for the lower side of the rectangular coil to pass through that position. The force on the coil is however the integral of $(I \times B)ds$. If the upper side were wholly in a region of zero flux density, if the lower side were in a region of constant flux density B, if the two long

sides were exactly vertical and separated by a known, well-defined, distance a, the force would just be BIa. Even the lower side of the coil however does not necessarily lie wholly in a region of constant B. The flux density must therefore be measured along the length of the side and differences allowed for. The top of the coil is in a region where the flux density, though admittedly much less than in the gap, is yet not zero, or even negligible, when the aim is an accuracy of 1 in 10^6. This flux density must thus be measured, also at a number of points along the side, and corrections applied. In order to overcome the great and real difficulty of finding the exact width a, coils of different widths have been used, or even coils widened by means of accurately known gauges. It is then possible to measure the difference of the forces and use the difference of width to calculate k' in eq (3). When the width is changed however, the four sides take up different positions in the gap, and as in practice the field is never strictly uniform, corrections must be applied.

The corrections are not easy to calculate even though the field may have been well "mapped." The conclusion seems to be that just as the current balance suffers from uncertainty in the dimensions of its coils and in the distribution of current in the wire, the determination of the field in the gap is beset by the difficulty of correcting the measured force for the variations of flux density and of coil shape. These corrections might well involve greater errors than those associated with the coils of a current balance.

The strong-field method on the other hand has this advantage that the force on the rectangular coil can be made larger than the force between the coils of a current balance, say 10 times or more, and the dead mass can be made smaller. The operation of weighing is thus facilitated and, possibly, the specification of the "magnetic balance" can be a little less stringent than that of a current balance. Protection from air currents and other disturbances is however as necessary for this system as it is for the current balance, and possibly even more difficult to achieve.

We must note too that although weighings by the current balance may be more laborious than by the magnetic balance, the measurements are planned to reduce the random error of weighing to a small fraction only of the systematic error. A similar procedure for the magnetic balance might admittedly be less lengthy and tiresome, but would not lessen the systematic errors associated with the coil and the mapping of the magnetic field. It is these errors which it may prove difficult to reduce to say 2 in 10^6.

I hope that this general survey will make it easier for members of the conference to follow the more specialized discussions on the program. I have not discussed particular papers nor included a list of references because in the first place it seemed to me that what you needed in the opening talk on the subject was general principles rather than details, and in the second place because of the excellent discussion and list of references already available in the scholarly work of Taylor, Parker, and Langenberg.[1]

DISCUSSION

C. H. Page: There is one other very sneaky error. Paul, that you didn't mention in the strong field method. That is making sure you know the orientation of B as well as its strength. Some of our experiments went awry because B wasn't exactly horizontal.

P. Vigoureux: Oh, yes. Yes.

J. Terrien: Just a comment, because it felt somewhat mysterious that by measuring the same constant by two methods you get the value of the true ampere. But the reason is clear I think. Because when you measure the ampere you have to measure the force between two coils. But the field is weak and the force is weak. In that case the gyromagnetic ratio of the proton is used to compare a weak field to a strong one and it enables you to measure the ratio of the fields.

P. Vigoureux: Yes.

J. Terrien: So that in the first method, the weak field, you produce a weak field. Then by means of the gyromagnetic ratio you can compare it to a strong field, and then the strong field you use to measure the force which is a stronger force. That's the reason when you eliminate the intermediate step of the gyromagnetic ratio you have exactly the current balance.

P. Vigoureux: I think that is very well put, sir, yes.

---

[1] The Fundamental Constants and Quantum Electrodynamics (Academic Press, New York/London, 1969).

Application of Nuclear Resonance to the Monitoring of Electrical Standards

R. L. Driscoll and P. T. Olsen

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Developments at Gaithersburg include compensation for earth's magnetic field variations and use of the method of nuclear induction to determine the precession frequency of protons in a magnetic field. Current procedures and recent results will be presented.

Key words: Ampere, earth's magnetic field, gyromagnetic ratio of proton, nuclear induction, nuclear resonance, ohm, solenoid, volt.

1. Introduction

There has been a progression in the accuracy of the electrical unit of current from the percent range 100 years ago to about 5 parts per million (ppm), which is about the best we can do in an absolute sense today. The precision with which we can now detect changes in a current by means of a current balance or rather $\mathcal{E}/R$, the ratio of emf over resistance used to control a current, is about 2 ppm. Resistance standards are relatively rugged things and can be checked independently in an absolute sense, and for this reason, the result of a current balance determination can be regarded, after putting in the independent value of the resistance standard, as assigning a value to the emf of the standard cell used to control the strength of the current. The current balance relation can be written $\mathcal{E}=R(F/F_0)^{1/2}$ where F is the observed force and F_0 is the balance constant calculated from linear dimensions of the balance coils.

Our involvement in the measurement of γ_p, the gyromagnetic ratio of the proton, brought about through the offices of Francis Bitter, E. U. Condon, R. D. Huntoon, and F. B. Silsbee, was for the purpose of getting a better value for γ_p rather than any notion that it might be used later to look for variations in the electrical standards.

Using the method of nuclear resonance absorption of Purcell, Torrey and Pound [1] and a magnetic balance similar to that of Briggs and Harper [2], Thomas Driscoll and Hipple [3] found $\gamma_p = 2.67523(6)\times10^8\ T^{-1}\ s^{-1}$ for protons in water when expressed in terms of NBS units maintained at the time. The number just given is too inconsistent with later measurements of γ_p and the ampere to be of more than historical interest.

Beginning about 1950 a solenoid to be used as part of a Pellat type electrodynamometer (PED) was constructed under the guidance of Charles Moon. It was kept in mind during the construction that this solenoid might be used to provide a magnetic field for a later measurement of γ_p. After completion of the solenoid the remainder of the PED was assembled and some observations were made with it before 1954. In the period (1954–1956) with the collaboration of R. D. Cutkosky, a final set of observations with the PED and a complete redetermination with the NBS current balance were carried out. On the basis of these measurements Cutkosky and Driscoll [4] found that the NBS ampere was about 10 ppm larger than the absolute ampere.

Arrangements were made for another cooperative effort with our atomic physics section under L. M. Branscomb. In the period 1954–1956 P. L. Bender designed equipment for the measurement of γ_p by the method of free precession which was appropriate for use in the solenoid from the electrodynamometer. R. D. Cutkosky collaborated on the construction and measurement of correcting windings for the solenoid [5] which provide a design field uniformity of $\sim 1:10^7$ over a 2-cm radius sphere at the center of the coils. Preliminary trials of the free precession method in the solenoid at Washington were not completely satisfying mainly because we had no provision for cancelling earth's magnetic field. After negotiation, the experiment was set up at the Fredericksburg Magnetic Observatory (FMO) where a large set of Helmholtz coils was made available to cancel the earth's field. Performance of the entire system at FMO met expectations and after some months of experience Bender and Driscoll [6] found $\gamma_p = 2.67515(1)\times10^8\ \mathrm{rad}\,T^{-1}\ s^{-1}$. When the facilities of FMO became unavailable, our equipment was returned to Washington. After checking the dimensions of the solenoid and constructing a 5 foot, 3 dimensional set of Helmholtz coils for cancelling earth's field, assembly of the apparatus and a set of observations was made with the collaboration of T. E. Wells. The published result [7] of this series $\gamma_p = 2.675156\times10^8\ \mathrm{rad}\,T^{-1}\ s^{-1}$ differs from the FMO result by about 3 ppm and is about as close as could be expected considering the difference of environments. The magnetic nature of the environment with

Table 1. γ_p *at Washington 1960–67, NBS units*

60–62	2.6751563
63	1545
64	1546
65	1554
66	1551
67	$2.675155_7 \times 10^8$ rad/T s

regard to local magnetic disturbances, gradients and magnetic materials in the neighborhood was poorer at Washington than it was at FMO. The possibility of realizing the assigned value of the standard cells was better at Washington than it was at FMO where the standard cell enclosure was exposed to wide changes in ambient temperature.

On account of the uniformity of results over a period of a year or so at Washington a note [8] was sent to BIPM in 1961 suggesting that with the specifications that the environment and apparatus remained unchanged, a change in the ratio $\mathcal{E}/R$ could be checked at any time to 1 ppm by the measurement of γ_p. Some measurements were made at Washington for several years following 1962 and using that series [7] as a reference the variations in table 1 were found.

Early measurements at Gaithersburg were made by the free precession method without prepolarization of the sample in a strong field as we did not want to contaminate our nonmagnetic building with a vacuum pump and a strong magnet. When operating without prepolarization a sample size of 37 mm dia was used and the signal to noise ratio

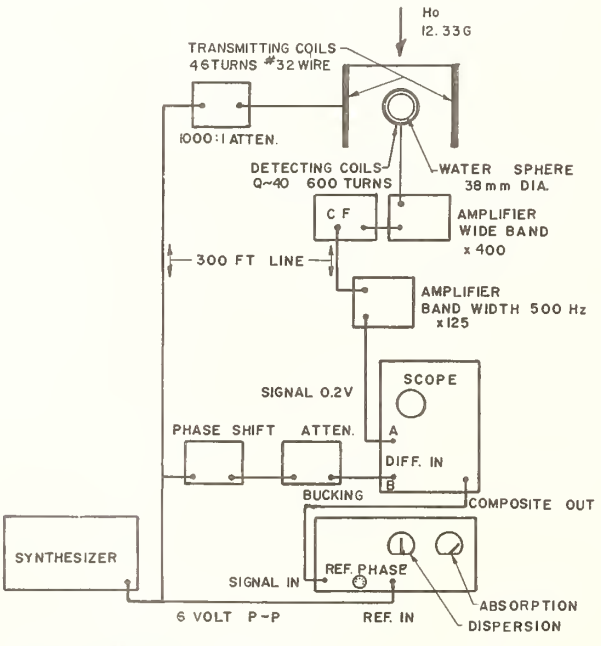


Figure 1. *Block diagram for measurement of precession frequency by method of nuclear induction.*

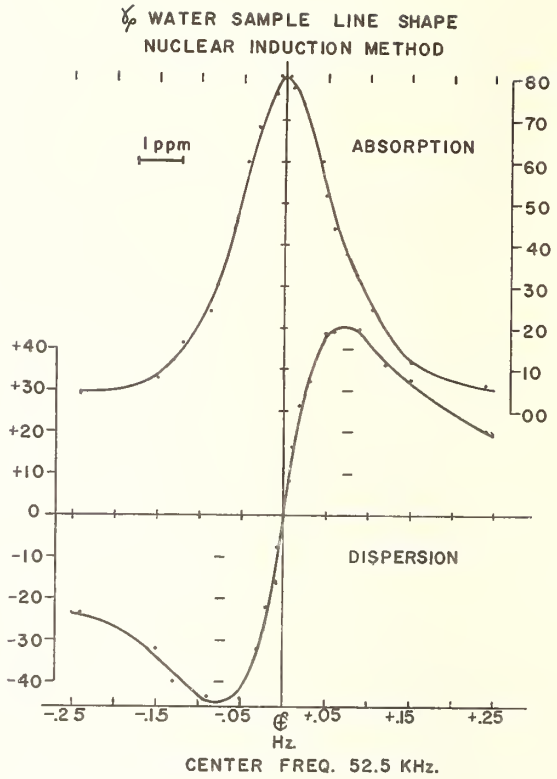


Figure 2. *Proton resonance line shape and separation of components.*

was ~3 to 1. Judging from the indicated frequency of artificial signals of known frequency which could be injected with a signal to noise ratio of 3 to 1, it was found that reliable measurements could be made without prepolarization.

In the summer of 1968 we began to measure the precession frequency of protons in a magnetic field by the method of nuclear induction using the arrangement shown in figure 1. This method provides a continuous signal whose frequency is taken as that of the driving source. It has been demonstrated that the frequency of the driving source when centered on the resonance line is essentially the same as that of the free precession signal as would be expected with a proton sample Q of $\sim 3 \times 10^5$. During early comparisons of the nuclear induction and free precession methods the statistics were poorer in the nuclear induction case on account of variations in the accuracy of setting to the center of the resonance line. All known difficulties of this kind with the nuclear induction method were gradually overcome to the degree that it is believed the frequency can be set to the center of the line to ~1 part in 10^7. The separation of the absorption and dispersion components of the resonance line displayed in figure 2 was plotted from readings of the phase detector indicators at fixed settings of the synthesizer.

Another factor contributing to the ease and reliability of our observations is the incorporation of means for eliminating earth's magnetic field varia-

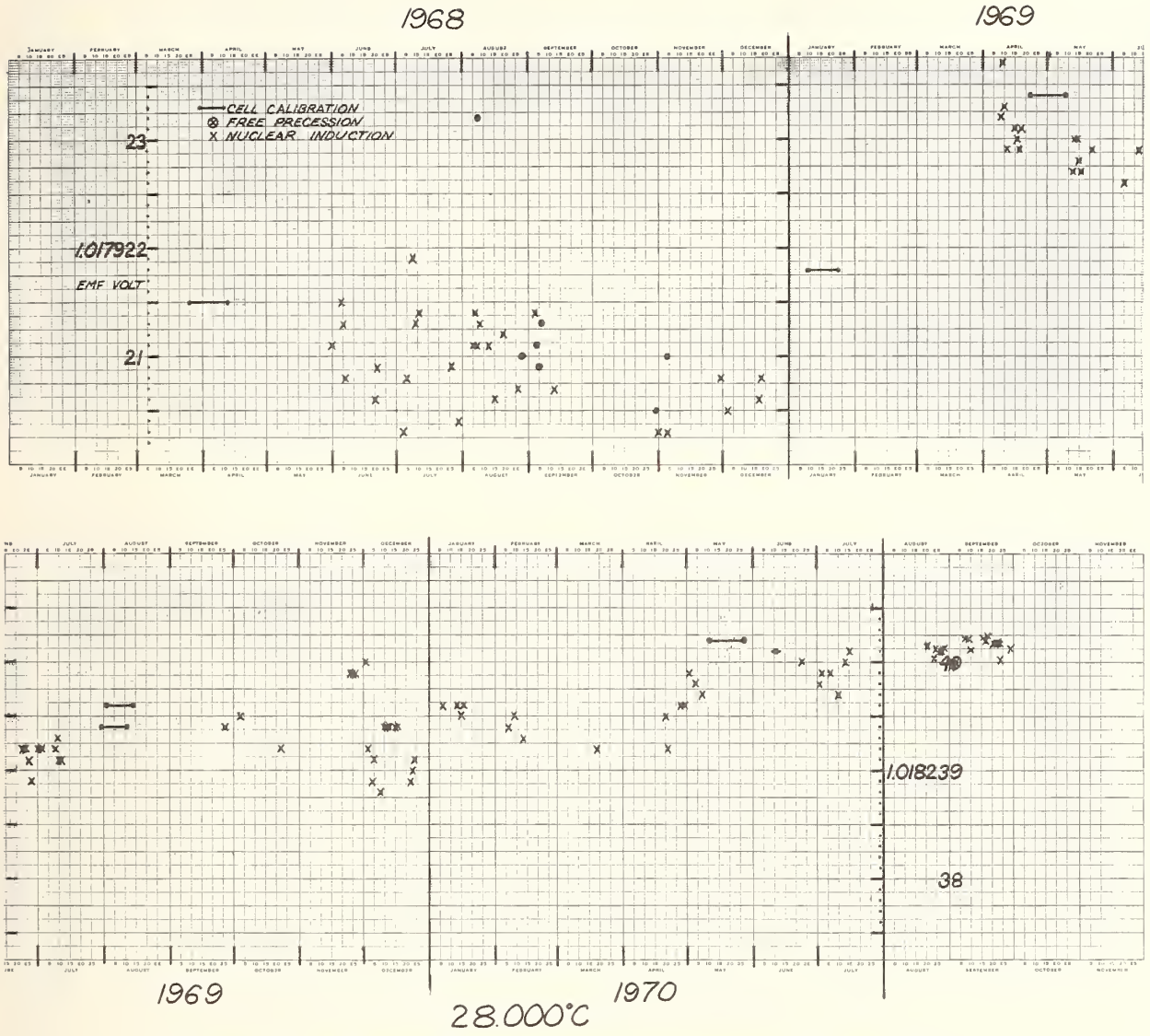


FIGURE 3. *Assigned and apparent mean emf of six standard cells in enclosure # 1461.*

tions along the solenoid direction. The earth's field compensating system, suggested by P. L. Bender, consists of a field controlled rubidium oscillator phase-locked to 10 kHz in a closed servo loop [9]. The correcting current from a lock-in amplifier is used to apply the same correcting field to the solenoid as it does to the oscillator 100 m distant, and to the extent that the variations are the same at both locations the compensation is exact. An exact figure for the degree of compensation cannot be given at this writing but over a 12 month period it has been found satisfactory for the intended purpose, that is, to maintain the stray field along the solenoid constant to a few tenths of a nanotesla over a 30 min period required for a set of observations.

2. Comparison of Apparent Values of EMF With Assigned Values

It is assumed that the assigned value R of the resistance standard is realized and that the solenoid dimensions remain constant and that any change in the observed precession frequency (after normalization for known interferences) can be ascribed to a change in the emf $\mathcal{E}$ of the standard cell. With precession frequency ν and magnetic field H, the relation with γ_p is

$$\gamma_p = 2\pi\nu/H$$

or

$$\gamma_p = 2\pi\nu R/K[1+\alpha(t-25)]\mathcal{E}$$

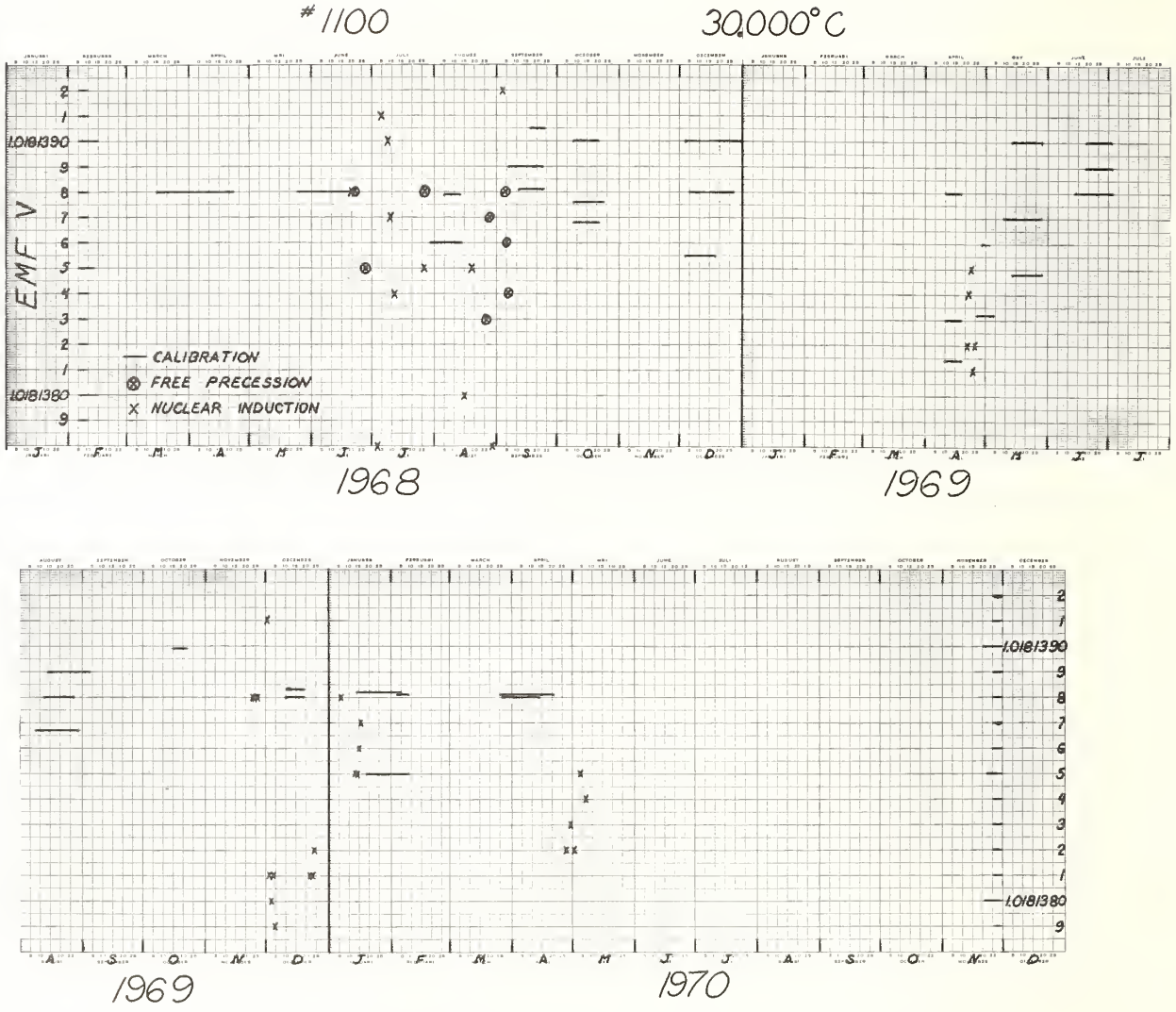


FIGURE 4. *Assigned and apparent emf of four cells in enclosure #1100.*

Where K is the field strength per unit current in the solenoid which was calculated from linear dimensions of the coil; and the constant $\alpha = -0.5 \times 10^{-6}/°C$ was determined experimentally by observing ν at various values of the solenoid temperature t while other things remained unchanged. After rearrangement the expression for the apparent emf is

$$\mathcal{E}_{app} = 2\pi\nu R/\gamma_p K[1+\alpha(t-25)].$$

For the purpose of looking for changes in $\mathcal{E}_{app}$ we adopt the γ_p value found here at Gaithersburg in 1968 which is with the 1969 change of units 2.6751301×10^8 rad T^{-1} s^{-1}. The data given in figure 3 has been gathered on a commercial temperature regulated enclosure containing six saturated cells. The same cells were used throughout and to make the chart meaningful the cell assignments as well as the apparent values have been normalized to a cell enclosure thermometer reading of 33.660 °C up to the point in time where the enclosure temperature was reduced to 28 °C. After the temperature change all values are normalized to 28.000 °C. Calibration assignments of the mean emf of the six cells in enclosure #1461 are indicated by solid lines; apparent values by the free precession method are indicated by circled crosses, and apparent values by the nuclear induction method are shown by crosses. The apparent values are dependent upon the actual mean emf of the six cells at any time of observation and may be expected to fluctuate in about the same manner as the mean fluctuates during a period of calibration. The scatter range of the mean was about 0.4 μV in the day to day normal calibrations which have been examined. Known uncertainties and their magnitudes as they affect the determination of the apparent values are (1) error in setting solenoid

current relative to $\mathcal{E}$ and R: 1 part in 10^7 (2) error in realization of solenoid temperature: 1 part in 10^7, and (3) error in setting frequency to center of resonance line: 1 part in 10^7. Item (3) could have been larger in some of the early points of 1968. The large change, in the calibration of February 1969, was due to intentional removal of accumulated gas from the negative limb of the cells. The separation in time between apparent values and calibration is due to the necessity of moving the enclosure to the standardizing laboratory. The results in figure 4 are on another enclosure identified as #1100. The variety in the concurrent calibrations resulted from a study of different instruments and methods used in the comparison of the cells with the primary group. Although the fluctuations in the apparent values have been larger than expected in consideration of the known uncertainties mentioned earlier, there has been no obvious increasing divergence between assigned and apparent values of emf during the interval of observations.

3. References

[1] Purcell, E. M., Torrey, H. C., and Pound, R. V., Phys. Rev. **55**, 526 (1939).
[2] Briggs, G. H., and Harper, A. F. A., J. Sci. Instruments **13**, 119 (1936).
[3] Thomas, H. A., Hipple, J. A., and Driscoll, R. L., J. Research NBS **44**, 569 (1950).
[4] Cutkosky, R. D., and Driscoll, R. L., J. Research NBS **60**, 297 (1958).
[5] Snow, C., and Driscoll, R. L., J. Research NBS **69C**, 49 (1965).
[6] Bender, P. L., and Driscoll, R. L., Phys. Rev. Letters **1**, 413 (1958).
[7] Driscoll. R. L., Phys. Rev. **136**, A54 (1964).
[8] Cutkosky, R. D., and Driscoll, R. L., Comité International des Poids et Mesures, 8e Session (1957).
[9] Olsen, P. T., and Driscoll, R. L., Compensation of Earth's Field Variations by Field Controlled Rubidium Oscillator. (In process).

DISCUSSION

H. Capptuller: We thank Mr. Driscoll very much for his paper, and I think we can't but admire the reproducibility of the results of his apparatus even after transportation from Fredericksburg to Gaithersburg.

J. S. Thomsen: I believe you remarked in your talk about the Thomas, Driscoll, and Hipple experiment. There has been comment in previous discussions of the constants about possible discrepancies through the field not being quite horizontal so that the protons looked at one field and the test coil measured another.

I just would like to say in that regard that, while one can't make any rigorous analysis, something like ten years ago I looked at the original paper and the field data that is given there, and, if one assumes that the field is normal to the pole faces and that Maxwell's equations apply in the air gap between the pole faces, you can at least make some estimates of how far the field can depart from horizontal. The result can't be called rigorous, but it indicates to me that it's highly unlikely—or at least unlikely—that the entire discrepancy could be due to the direction of the field.

R. Driscoll: The difference between that Thomas, Driscoll, Hipple work and low field work as related to the ampere shows that the Thomas, Driscoll, Hipple declared uncertainty—that the truth might be two or three times the probable error. And I don't know that there's anything sacred about repeating and agreeing within the stated probable error. (*Laughter*)

J. Terrien: When you use a constant dimensional coil and γ_p to check the constancy of the standards which maintain the ampere, is it necessary to check the geometrical dimensions of the coil and the change during the years? And with what uncertainty is that being done?

R. Driscoll: Thank you for bringing out that point which is specified in our paper but lack of time prevented me from reading. The pitch of the coil is of first order importance in the measurement of γ_p but constancy is all that is required for the present purpose. The coil has been measured a number of times since its construction in 1950. Measured values of the pitch have fluctuated over a total range of about 4 ppm and I believe that difficulties in the measuring process have more to do with the observed fluctuations than actual variations in the pitch with time. When checked within the past year a value intermediate in the 4 ppm range was observed. With the advent of the laser interferometer it may now be possible to determine the mean pitch and its local variations with more accuracy than has been realized up to this time.

A γ_p' Determination at the ETL; Two Proposals for Obtaining a Uniform Calculable Magnetic Field for an Atomic Standard of Current

Ko Hara and Hisao Nakamura

Electrotechnical Laboratory 5-4-1, Mukodai, Tanashi, Tokyo, Japan

Two methods for obtaining improved calculable magnetic fields for low field γ_p' determinations are discussed. In the first, a single layer solenoid of proper proportion is modified to a solenoid of fourth order uniformity, without any correction coil for uniformity. In the second, a solenoid is confined in a box of high μ material. This gives a calculable uniform field because of imaging effects. A practical arrangement is given for use when μ is not extremely high. In this method, isolation from external fields is achieved.

Key words: Calculable uniform magnetic field; γ_p' determination.

1. Introduction

Table 1 shows the error budget of the γ_p' determination at the ETL in 1968, which shows that uncertainties due to the solenoid system are predominant. Although the solenoid used in this particular determination has a relatively small length to diameter ratio of unity and is not well suited for the purpose, it may generally be said the heart of a γ_p' determination consists in having a good calculable magnetic field which is as accurate, uniform and stable as possible. Difficulties which are encountered in a low-field γ_p' determination can be stated as follows: (1) A correction coil for field uniformity is a source of troubles both in its fabrication and its alignment with the main solenoid, and furthermore it gives a relatively small volume of uniform field. (2) A small volume of uniform field necessitates polarization of the water sample. This leads to heat generation and complexity around the water sample which introduces sources of errors. (3) To be calculable, a solenoid is supposed to be placed in a nonmagnetic environment. This is rarely realized because of fluctuation of the geomagnetic field of natural and human origins.

Table 1. *Error estimation of the γ_p' determination at the ETL in 1968*

Source of error		Estimated error
		ppm
Field constant	Diameter	0.7
	Axial length	0.7
	Diameter of winding	0.4
	Temperature	0.3
	Current distribution	1.4
	Field coil alignment	1.0
	Diameter instability	2.0
Current	Resistance calibration	0.3
	Standard cell calibration	0.3
	Current measurement	0.5
	Temperature	0.8
Period	Frequency	0.1
	Period determination	0.3
	Fluctuation of earth's field	1.0
	B_e measurement	0.1
RMS total		4

Table 2. *Important expansion coefficients of a field by a helical winding pair*

$$Z_0^0 = 2q_1$$
$$Z_2^0 = (3/2)q_5$$
$$Z_1^1 = (-1)^N (p/a)6(q_5 - 5q_7)\sin\theta$$
$$Z_0^2 = -3q_5$$
$$Z_4^0 = -(15/8)[q_7 - (7/4)q_9]$$
$$Z_2^2 = 15[q_7 - (7/4)q_9]$$
$$Z_0^4 = -5[q_7 - (7/4)q_9]$$
$$Z_6^0 = (5/16)[q_9 - (9/2)q_{11} + (33/8)q_{13}]$$
$$Z_4^2 = -(315/8)[q_9 - (9/2)q_{11} + (33/8)q_{13}]$$
$$Z_2^4 = (105/2)[q_9 - (9/2)q_{11} + (33/8)q_{13}]$$

In this paper two design principles for a calculable solenoid are introduced which may overcome some of these difficulties. Before going into details, field formulae for a pair of helical windings (as shown in figure 1) must be introduced [1]. The center field due to a helical winding pair is given in cylindrical coordinates as

$$H_z = \frac{2\pi I}{g} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} Z_n^m(\theta) \left(\frac{Z}{a}\right)^m \left(\frac{\rho}{a}\right)^n \qquad (1)$$

where the coefficients Z_n^m are given in table 2. The

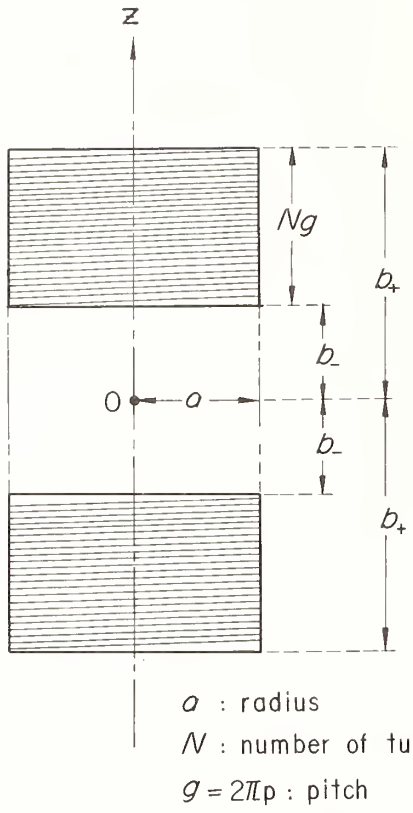


FIGURE 1. *A symmetric pair of helical windings.*

q_n in table 2 are

$$q_n = \frac{b_+/a}{[1+(b_+/a)^2]^{n/2}} - \frac{b_-/a}{[1+(b_-/a)^2]^{n/2}} \qquad (2)$$

2. A Modified Single Layer Solenoid

Figure 2 shows an outline of a modified single layer solenoid which has a pair of stepped (shorted) turns of axial length $2k$ at an axial distance b from the center. This is understood as a combination of a main solenoid and a correcting coil pair at the stepped position with current of equal magnitude and opposite polarity of that of the main solenoid. When the values of a, b, n, and k are properly adjusted, the

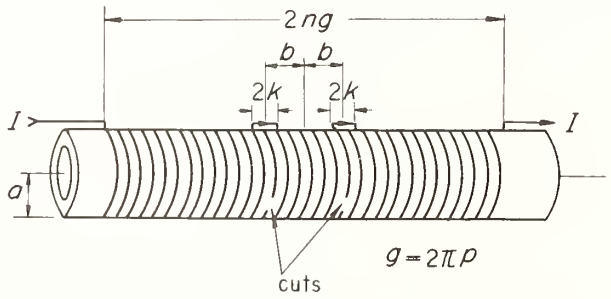


FIGURE 2. *Parameters of a modified single layer solenoid.*

TABLE 3. *Examples of modified single layer solenoid*

n	a (cm)	b (cm)
412	10.9233	3.35
404	10.7534	3.30
396	10.5831	3.25
388	10.4124	3.20
381	10.2758	3.15

$g = 2k = 0.1$ cm.

center field can be expressed as

$$\frac{g}{2\pi I} H_z = Z_0^0 + Z_1^1 \left(\frac{z}{a}\right)\left(\frac{\rho}{a}\right) + Z_4^2 \left(\frac{z}{a}\right)^2\left(\frac{\rho}{a}\right)^4 + Z_2^4 \left(\frac{z}{a}\right)^4\left(\frac{\rho}{a}\right)^2 + Z_0^6 \left(\frac{z}{a}\right)^6 + F(\rho, z, \theta). \qquad (3)$$

Examples of these parameters are given in table 3 and values of the Z_n^m are given in table 4. The residual term F is negligibly small. In table 5, limits of tolerances of dimensional measurement and machining accuracy are given for 1 ppm computed accuracy of the center field and 1 ppm uniformity in a sphere of 4 cm diameter.

This modified single layer solenoid has the feature that it gives a uniform field from a single solenoid, so that any adjustment of a separate correction coil is unnecessary. The volume of uniform field is just the same as that of the single layer solenoid with a correction coil which is widely used at present.

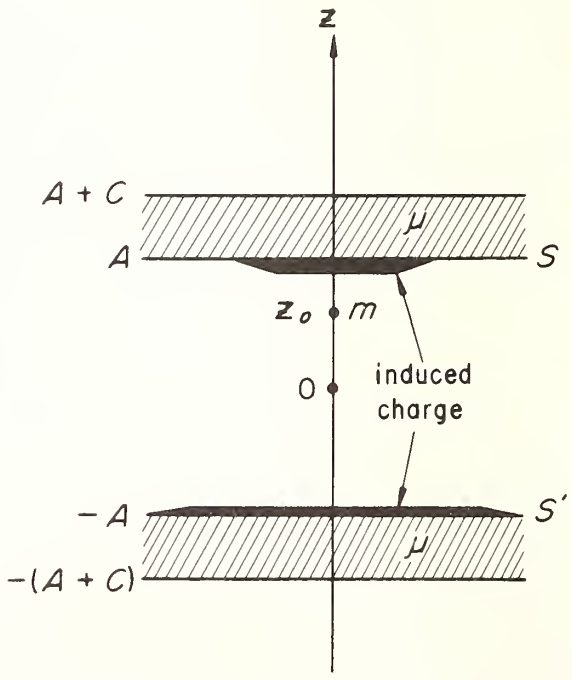


FIGURE 3. *A point charge between a plane parallel slab pair.*

TABLE 4. *Coefficients of main and inhomogeneous terms*

n	Z_0^0/p	Z_1^1/pa^2	Z_4^2/pa^6	Z_2^4/pa^6	Z_0^6/pa^6
412	120.46178	$-2.92\sin\theta\times10^{-6}$	-3.6×10^{-6}	4.8×10^{-6}	-0.6×10^{-6}
404	120.41458	$5.84\sin\theta\times10^{-6}$	-4.0×10^{-6}	5.4×10^{-6}	-0.7×10^{-6}
396	120.36576	$-3.30\sin\theta\times10^{-5}$	-4.5×10^{-6}	6.0×10^{-6}	-0.8×10^{-6}
388	120.31526	$6.59\sin\theta\times10^{-6}$	-5.1×10^{-6}	6.7×10^{-6}	-0.9×10^{-6}
381	120.25960	$-7.00\sin\theta\times10^{-6}$	-5.5×10^{-6}	7.4×10^{-6}	-1.0×10^{-6}

3. A Magnetically Isolated Calculable Solenoid (MICS)

It is natural to imagine that a solenoid surrounded by high permeability material is not calculable at all. The magnetically isolated calculable solenoid (MICS) is, however, a calculable example of such a device. We begin by considering the field due to a point magnetic charge surrounded by a rectangular box of high permeability material. This is because a field due to a steady current is equivalent to that due to a magnetic double layer of a corresponding geometry. The field between S and S' due to a point charge m at z_0 between a pair of plane parallel slabs of high permeability material (see fig. 3) is equivalent to the field due to the true charge plus an infinite series of point charges which represent the effect of the induced charges on the boundaries of the slabs and are understood to be images of the true charge. Once a solution of the type indicated in figure 3 is obtained, the field due to a charge surrounded by any number of pairs of plane parallel boundaries can be obtained in principle by successive application of the imaging process.

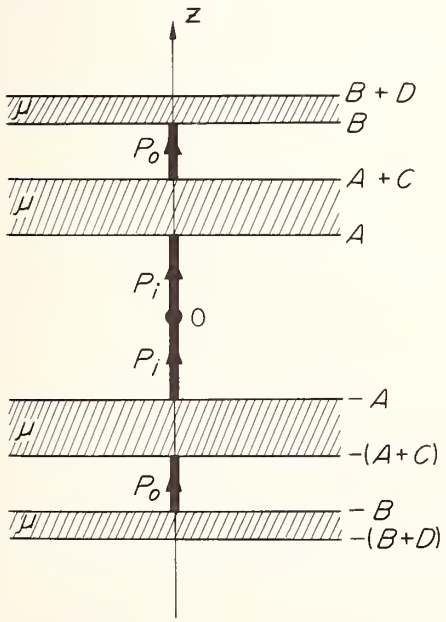


FIGURE 4. *Dipole line pairs between double slab pair.*

Thus the field due to a current distribution surrounded by a box is soluble.

The magnetic potential V produced by induced charge distributed over the boundaries of double pairs of plane parallel slabs of relative permeability μ due to a dipole line of strength p_i and a dipole line pair of strength p_0 is expressed at a point (ρ, z) (see fig. 4)

$$V=\int_0^\infty F(e^{\lambda z}-e^{-\lambda z})J_0(\lambda\rho)d\lambda, \qquad (-A<z<A) \quad (4)$$

where

$$F=F_1+F_i+F_0 \quad (5)$$

$$F_1=p_i\frac{Mx}{1-Mx}(e^{-\lambda A}-e^{\lambda A}) \quad (6)$$

$$F_i=-p_i(1-M^2)\frac{xy}{(1-M^2y)(1-Mx)^2}(e^{-\lambda A}-e^{\lambda A}) \quad (7)$$

$$F_0=p_0(1-M^2)\frac{xy^{1/2}}{(1-M^2y)(1-Mx)^2}(e^{-\lambda A}-e^{\lambda A}) \quad (8)$$

$$M=(\mu-1)/(\mu+1), \qquad (1-M^2)=4\mu/(\mu+1)^2, \quad (9)$$

$$x=e^{-2\lambda A}, \qquad y=e^{-2\lambda C}, \quad (10)$$

and J_0 is the zeroth order Bessel function. In these expressions, terms of order equal to or higher than $(1-M^2)^2$ are neglected, whereby $\mu\gg1$ is assumed. F_1 gives the term due to p_i when the thickness C of the inner slab pair tends to infinity. F_i gives the correction term due to p_i in case of finite thickness. F_0 is the term due to p_0. The effect of the outer slab pair is included in terms of order of $(1-M^2)^2$ or

TABLE 5. *Limit of tolerances of dimensional measurement and machining accuracy in case of examples of table 3*

	measurement	machining
dg	-1.07×10^{-7} cm	-1.25×10^{-4} cm
dn	6.22×10^{-3} turn	-5.13×10^{-1} turn
2dk	-1.20×10^{-5} cm	-4.57×10^{-4} cm
db	1.56×10^{-3} cm	8.47×10^{-3} cm
da	-1.82×10^{-4} cm	1.74×10^{-2} cm

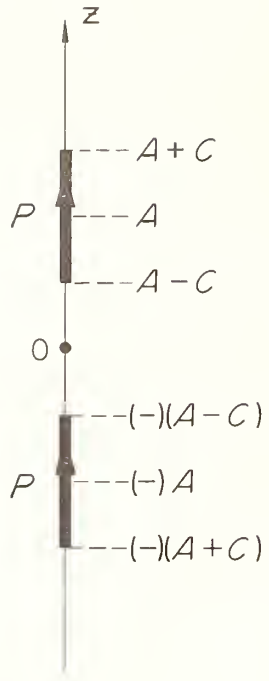


FIGURE 5. *A dipole line pair.*

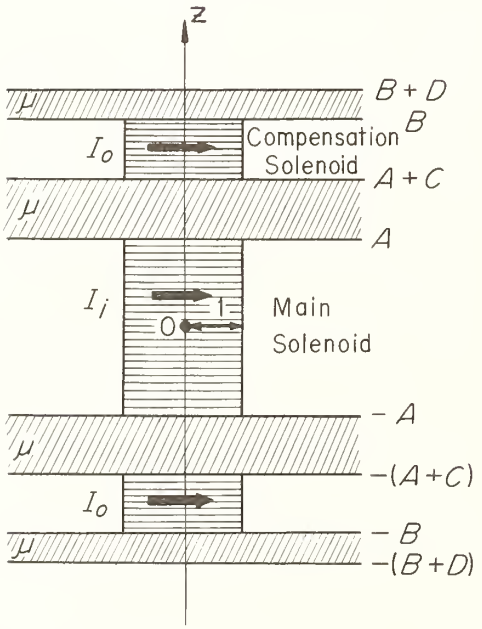
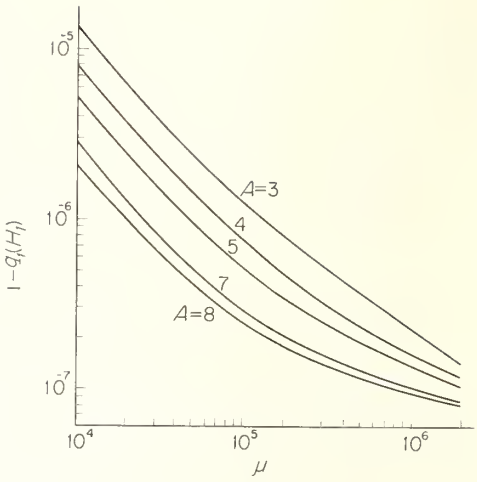


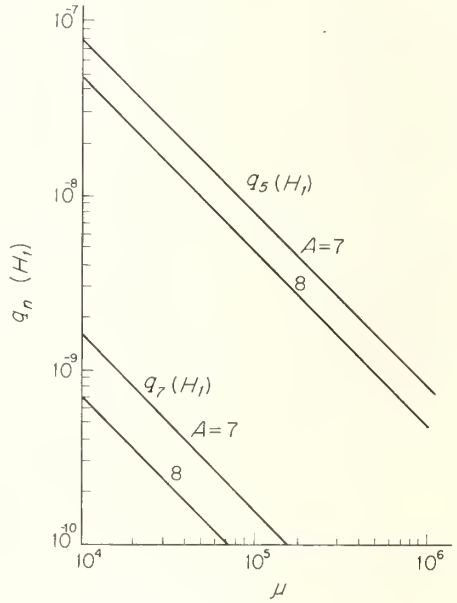
FIGURE 6. *Solenoid pairs between double slab pairs.*

TABLE 6. $q_n(H_i)$ ($\mu = 2\times10^4$, $A = 8$, $C = 0.2$)

	q_1	q_5	q_7
H_1	$1-1.03\times10^{-6}$	-2.4×10^{-8}	-3.6×10^{-10}
H_2	9.7×10^{-7}	-2.5×10^{-8}	-4×10^{-10}
H_0	1.44×10^{-4}	-1.5×10^{-7}	-5×10^{-10}



(a) $1-q_1(H_1)$.



(b) $q_5(H_1)$ and $q_7(H_1)$.

FIGURE 7. $q_n(H_1)$ *dependence on* μ *and* A.

higher. When $p_i = p_0 = p$, F becomes F_+ where

$$F_+ = F_1 + F_2 \tag{11}$$

$$F_1 = p \sum_{n=1}^{\infty} (Mx)^n (e^{-\lambda A} - e^{\lambda A}) \tag{12}$$

$$F_2 = p(1-M^2) \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (M^2y)^m (Mx)^n x^{1/2} y^{3/4} \times (e^{-\lambda C/2} - e^{\lambda C/2}). \tag{13}$$

When $p_i = -p_0 = p$, F becomes F_- where

$$F_- = F_+ - 2F_0 \tag{14}$$

$$F_0 = p(1-M^2) \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (n+1)(M^2y)^m (Mx)^n xy^{1/2} \times (e^{-\lambda A} - e^{\lambda A}). \tag{15}$$

Keeping in mind that the potential due to the dipole line pair of strength p shown in figure 5 is

$$V=p\int_0^\infty e^{-\lambda A}(e^{-\lambda C}-e^{\lambda C})(e^{\lambda z}-e^{-\lambda z})J_0(\lambda\rho)d\lambda$$
$$[-(A-C)<z<(A-C)], \quad (16)$$

the potentials given by eqs (12), (13), and (15) are understood to be the potentials due to superposition of infinite series of dipole line pairs.

The results for a dipole line are easily translated into the case of the solenoid shown in figure 6. Using a notation $[I,\ X\rightarrow Y]$ as a solenoid pair carrying a current I and extending from X to Y and from $-Y$ to $-X$, image solenoid pairs corresponding to F_1, F_2, and F_0 are as follows.

F_1:

$$[IM^{n+2},\ (2n+1)A\rightarrow(2n+3)A] \qquad (n=0,\ 1,\ 2\cdots) \quad (17)$$

F_2:

$$[I(1-M^2)M^{2m+n},\ (2n+1)A+(2m+1)C \rightarrow(2n+1)A+(2m+2)C] \qquad (m,\ n=0,\ 1,\ 2\cdots) \quad (18)$$

F_0:

$$[I(1-M^2)M^{2m+n}(n+1),\ (2n+1)A+(2m+1)C \rightarrow(2n+3)A+(2m+1)C] \qquad (m,\ n=0,\ 1,\ 2\cdots). \quad (19)$$

Here and hereafter axial dimensions such as A or C are normalized to the radius a. We shall denote the z-component of magnetic field due to $[I,\ (-A)\rightarrow A]$ and the sum of eq (17) as H_1. In the same manner, the field due to the sum of eq (18) is denoted as H_2 and the field due to the sum of eq (19) is denoted as H_0. It may be seen that H_1 is equal to the field of an infinitely long solenoid and H_2 and H_0 tend to zero when μ tends to infinity. H_2 and H_0 are small correction terms and their approximated expressions are as follows, where the notation $q_n(H_i)$ means a q_n of H_i field.

$$q_1(H_2)\doteqdot\frac{1-M^2}{4A^2}\sum_{i=0}^{\infty}\frac{1}{(2i+1)^2} \quad (20)$$

$$q_n(H_2)\doteqdot-\frac{1-M^2}{2A^{n-1}}\sum_{i=0}^{\infty}\frac{1}{(2i+1)^{n-1}} \qquad (n\geqq 5) \quad (21)$$

$$q_1(H_0)\doteqdot-\frac{1-M}{4AC}\log(1-M) \quad (22)$$

$$q_n(H_0)\doteqdot-\frac{1-M^2}{2^{n-1}A^{n-2}C}\sum_{i=1}^{\infty}\frac{1}{i^{n-2}} \qquad (n\geqq 5). \quad (23)$$

Values of $q_n(H_1)$ are computed using eq (1) as

$$q_n(H_1)=\frac{A}{[(1+A^2)^{1/2}]^n}+\sum_{i=1}^{\infty}M^i \times\left[\frac{(2i+1)A}{\{[1+(2i+1)^2A^2]^{1/2}\}^n}-\frac{(2i-1)A}{\{[1+(2i-1)^2A^2]^{1/2}\}^n}\right]. \quad (24)$$

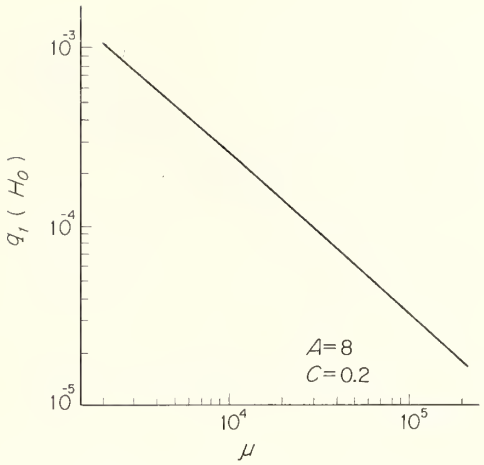


FIGURE 8. *μ dependence of $q_1(H_0)$.*

Dependences of $q_n(H_1)$ on μ and A are shown in figure 7. Numerical values of $q_n(H_i)$ are given in table 6 when $A=8$, $C=0.2$ and $\mu=2\times10^4$. From these, it may be concluded that the field H_1+H_2 around the center of the solenoid is uniform within 1 ppm in a volume of $(\rho/a)<1$ and $|z/a|<1$. It is necessary to know the value of μ when one wants to evaluate $q_1(H_1)$ and $q_1(H_2)$ with an accuracy better than 1 ppm. The difference $[q_1(H_1)+q_1(H_2)]-[q_1(H_1)+q_1(H_2)-2q_1(H_0)]=2q_1(H_0)$ can be determined experimentally with reasonable accuracy if the difference in proton precession frequencies for positive and negative polarities of the current through the compensation solenoid shown in figure 6 and the approximate value of γ_p' are known. The dependence of $q_1(H_0)$ on μ when $A=8$ and $C=0.2$ is shown in figure 8, and thus the value of μ can be found by comparing the measured value of $q_1(H_0)$ with figure 8. Therefore, the center field of the solenoid when the currents of the main and the compensation solenoids are common in magnitude and polarity is calculable with an accuracy better than 1 ppm, provided the magnetization curve of the slab material is linear and single valued to a sufficient degree.

B–H curves of a commercial supermalloy of 1.2 cm^2 cross section are shown in figure 9 with three different tracing periods. This test sample is so thick that hysteresis due to eddy current loss is appreciable, so that the d-c hysteresis curve has not been obtained even with the longest period available. Figure 10 is a plot of the apparent coercive force h_c versus the tracing period. One may thus assume that h_c at dc is of the order of 0.01 mOe. Taking the applied field of 10 Oe into account, the hysteresis of this sample might not cause ambiguity much greater than 1 ppm.

Sections free of windings are necessary at both ends of the main solenoid in order to introduce cables into the center part of the solenoid and also to handle the solenoid. For this purpose, we shall consider the geometry shown in figure 11, in which a

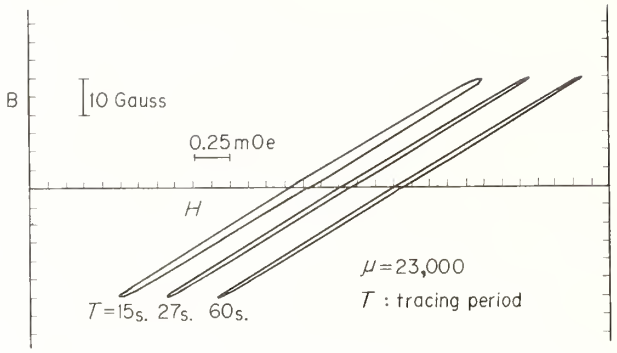


FIGURE 9. *B–H curve of a thick supermalloy sample at low magnetization.*

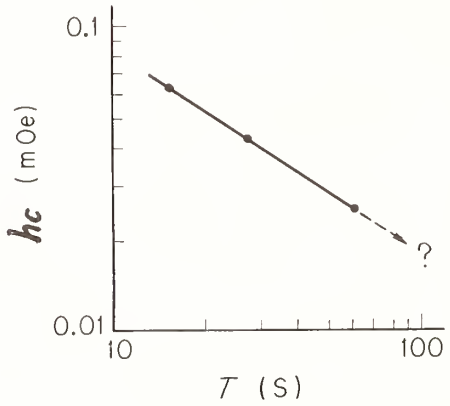


FIGURE 10. *h_c versus tracing period.*

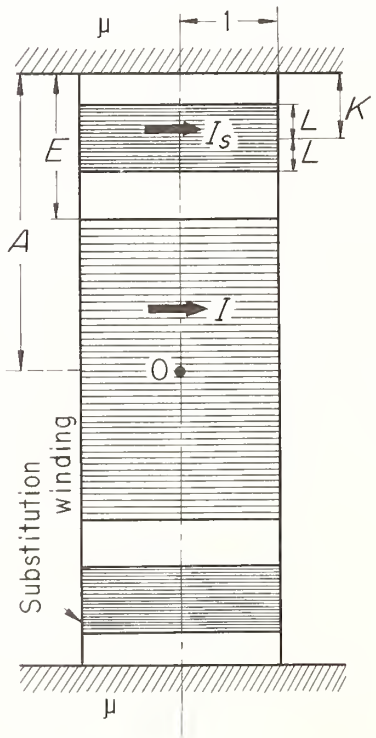


FIGURE 11. *MICS with a substitution winding.*

winding pair between (A−E) and A is replaced by a pair of windings between (A−K−L) and (A−K+L). The change of q_n of the main solenoid by this substitution is denoted as dq_n. When $\mu=\infty$, dq_n is found to be

$$dq_n=2\sum_{j=0}^{\infty}\sum_{k=\text{odd}}^{\infty}\frac{y_n^{(k)}\{x=(2j+1)A\}}{k!}\times\left[\frac{I_s}{I}\{(K+L)^k-(K-L)^k\}-E^k\right] \quad (25)$$

where

$$y_n=x/(1+x^2)^{n/2} \quad (26)$$

and

$$y_n^{(k)}=d^k y_n/dx^k. \quad (27)$$

When the conditions

$$IE=2I_sL \quad (28)$$

$$E^2=3K^2+L^2 \quad (29)$$

are satisfied, dq_n is zero to the third power of a small quantity. Equation (28) is just the condition for ampere-turn compensation. When μ is finite, q_n is computed numerically. Figure 12 shows the dependence of dq_n on A, E, μ, and L under the conditions of eqs (28) and (29). It is found that dq_n is almost independent of μ. When $A=8$, E as large as 1 is allowed without introducing any loss in uniformity and the center field is calculable with an accuracy better than 1 ppm.

Until now, the effect of side walls has not been taken into account. A MICS necessarily assumes confinement of the main solenoid in double rectangular boxes. Charges are then induced on the side walls of the boxes and these make unwanted images which modify the center field. To evaluate this effect, two assumptions have been made. (1) Permeability is infinite. (2) The wall thickness is infinite. The first assumption is equivalent to assuming that unwanted images result from leakage flux through windingless gaps which are introduced by employing a substitution winding pair. The second is based on the fact that the effect of the side walls is of the order of $1/\mu$ and the effect of the wall thickness is also of the order of $1/\mu$.

After some mathematics, the resultant field δH_z due to the side walls at the center is found to be

$$\delta H_z\doteqdot-(15/16)\pi IE^5(Y_0-\tfrac{1}{2}Y_{0.5}) \quad (30)$$

where

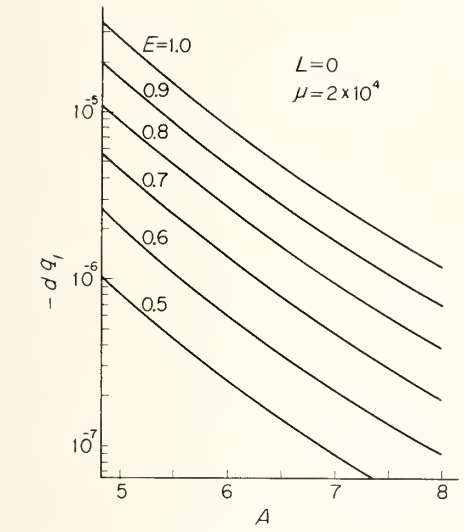
$$Y_n=\frac{16}{X_n^7}-8\frac{(7+W_n^2)}{X_n^9}+42\frac{(1+W_n^2)(1+3W_n^2)}{X_n^{11}}-63\frac{W_n^2(1+W_n^2)^2}{X_n^{13}} \quad (31)$$

and

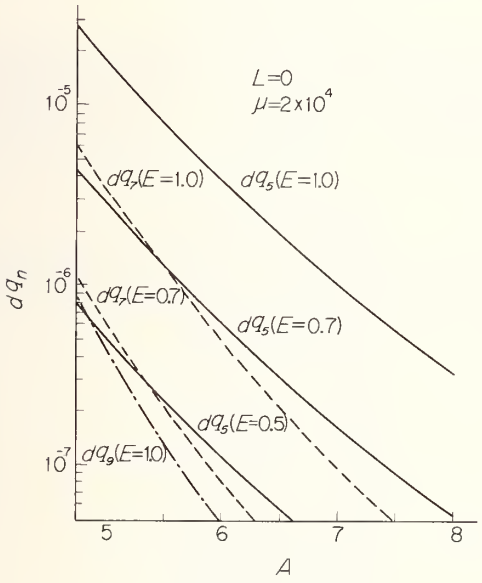
$$W_n^2=[(0.5)^2+n^2]W^2 \quad (32)$$

$$X_n^2=1+A^2+W_n^2. \quad (33)$$

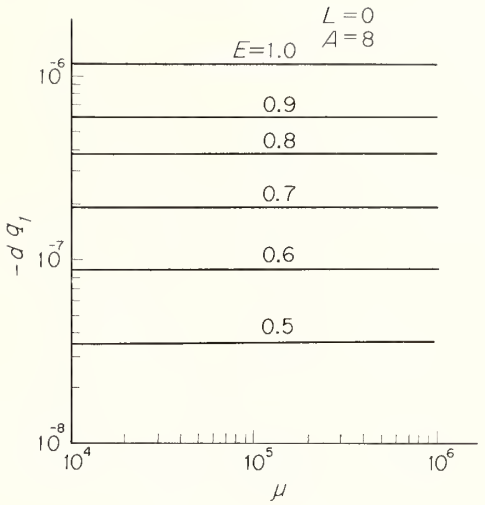
W is the separation between the side walls. When $A=8$, $E=1$ and $W=4$, $\delta H_z=-3\pi I\times10^{-6}$ which is about −0.8 ppm of the main field. Thus the effect due to unwanted images caused by the side walls is almost negligible.



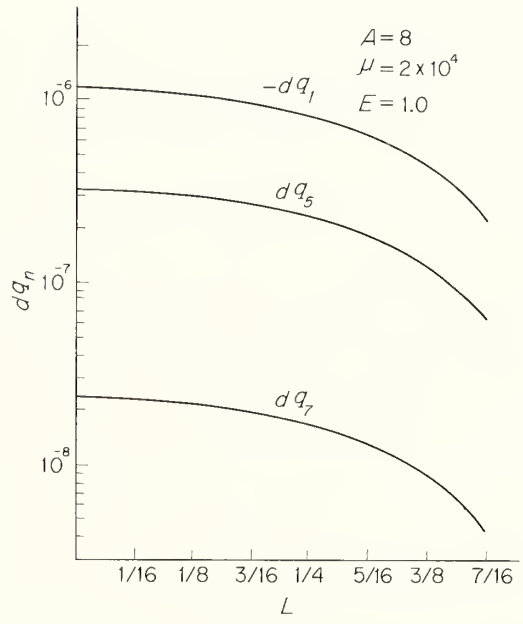
(a) dq_1 versus A and E.



(b) dq_n versus A and E.



(c) dq_1 versus μ and E.



(d) dq_n versus L.

FIGURE 12. *dq_n versus A, E, μ, and L.*

Current through the lead wires into the main solenoid also produces a field around the center. Assuming the geometry shown in figure 13, this is evaluated to be

$$\frac{(H_z)_l}{H_1} = \left(\frac{15}{8\pi}\right) gl^2WAz\rho$$

$$\times\{[(\rho^2+W^2+4A^2)^7]^{-1/2}-[(2W^2+4A^2)^7]^{-1/2}\} \quad (34)$$

where g is the pitch of the main solenoid, and z and ρ are cylindrical coordinates around the center. The ratio given by eq (34) is far less than 1 ppm, when $g=0.02$, $l=1$, $W=4$, $A=8$, and $|z|<1$, $\rho<1$. This evaluation is based on collecting the effect of all images due to six walls of the box.

In summary, we have encountered no serious objection to the possibility of a MICS for producing a calculable field with an accuracy of about 1 ppm. We have no experimental evidence of our own which confirms that the theory of a MICS is indeed correct to the order of 1 ppm. There is, however, an experimental study of modification of the magnetic field distribution between the pole faces of a cyclotron to obtain a suitable field profile by means of coils placed between the pole faces [2]. The theory in this case is also based on the summation of pole face image effects and fairly good agreement is obtained with experiment.

Considering that the heart of a low-field γ_p' determination is the necessary calculable solenoid,

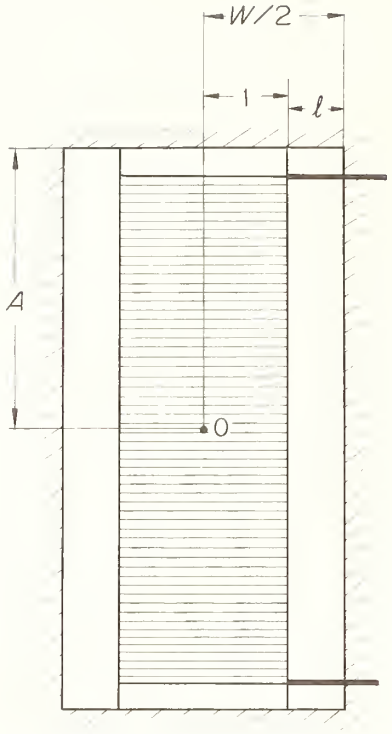


FIGURE 13. *MICS with a pair of leads.*

we have presented two new design principles for a calculable solenoid. A modified single layer solenoid provides a small step of improvement in giving an uniform calculable field without requiring any correction coil for field uniformity. About 0.1 ppm accuracy may be possible, if the finest machining and dimensional measurement are available. A nonmagnetic environment is, however, necessary as before.

In contrast, a MICS is isolated from external magnetic fields to a high degree and thus a nonmagnetic environment is not necessary at all. Unfortunately, the accuracy reached by a MICS depends on the magnetic behavior of the isolating boxes. This is hard to include in the theory in a complete form. In spite of this difficulty, we are going to build a MICS starting this year. This is because we have difficulties in finding a nonmagnetic environment on the one hand and we assume the ambiguity coming from the magnetic property will be removed by experimental procedure on the other.

4. References

[1] Hara, K., Denkishikenjo Kenkyuhokoku, No. 560 (1957) (Researches of the Electrotechnical Laboratory, No. 560 (1957)).

[2] Nakanishi, N., et al, Japanese J. of Appl. Phys., **8**, 1229 (1969).

DISCUSSION

H. CAPPTULLER: Dr. Ko Hara's paper shows us the difficulties in determining the field inside a single-layer solenoid, in the center of the solenoid. We have seen some complicated equations here for evaluating the field inside. Of course, by means of modern electronic computers this evaluation takes not much time, but the geometric measurement of the dimensions of the coil takes most of the time by our experiences.

The Gyromagnetic Ratio of the Proton Measured in a High Magnetic Field

B. P. Kibble and G. J. Hunt

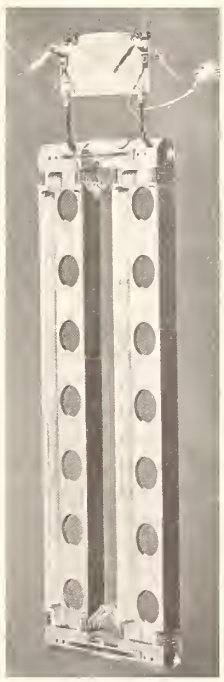
National Physical Laboratory, Teddington, Middlesex, England

We report measurements, using a prototype apparatus, of the force between a constant current in a coil with parallel vertical sides whose separation is changed by a known amount and a uniform horizontal magnetic field applied to the lower end. Simultaneous observation of the precessional frequency of protons in water in the field yields the proton gyromagnetic ratio in the maintained NPL unit of current. Combining the result with that obtained by observing the precessional frequency in the field of a precision solenoid carrying a related current will give the currents and gyromagnetic ratio in absolute units.

Key words: Ampere; Cotton balance; gyromagnetic ratio; proton.

1. Introduction

The reproducibility and constancy of the unit of current as maintained by the various national laboratories is likely to be significantly improved in the near future with the advent of the ac Josephson effect and the calculable capacitor, so there is a need for an improvement in the determination of the ampere in the fundamental units of mass, length and time. Also many experiments on fundamental constants involve measurements of magnetic fields in terms of the precessional frequency of protons in water. It is therefore important to obtain the proton gyromagnetic ratio γ_p' in water in absolute units to as high an accuracy as possible. Combination of two different methods of measuring γ_p' can solve both problems simultaneously, if we assume the value of γ_p' is independent of magnetic field.



The plate

At the NPL Vigoureux has already performed very accurate current balance and low field γ_p' experiments [1, 2], and so it is desirable to perform a high field experiment to a similar accuracy. We have conducted a feasibility study to find the difficulties likely to be encountered, and we report the results obtained with a prototype apparatus assembled for the purpose.

The high field method involves measuring the vertical force F_y on a coil carrying a current i_H with parallel vertical sides separated by a distance l in the x–y plane with the lower end in a uniform high magnetic field B_z whilst simultaneously observing the precessional frequency ω_H of protons in a sample of water. If the force is determined by balancing it with a mass M and the acceleration due to gravity is g then

$$F_y = Mg = B_z i_H l = (\omega_H/\gamma_p')\, l i_H.$$

Hence γ_p' can be measured in terms of M, g, ω_H, l and i_H.

2. The Apparatus

This is based on that used by Cappuller [3, 4] in which the width l of the coil is varied to eliminate many serious systematic errors. The rectangular coil is shown in the plate, and in figure 1. The vertical wires A, of 0.75 mm diameter Lewmex M grade insulated copper are cemented with 'Araldite' epoxy resin to the faces of girders machined from magnesium aluminium alloy, chosen for its low density and magnetic susceptibility, and high Young's modulus. Alternative sets of wires not used in the work reported here were cemented to the outer faces. The

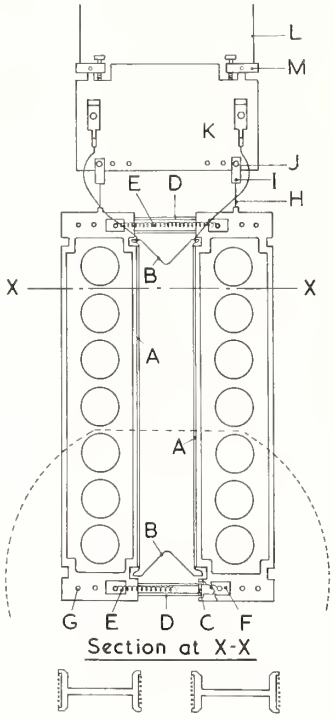


FIGURE 1

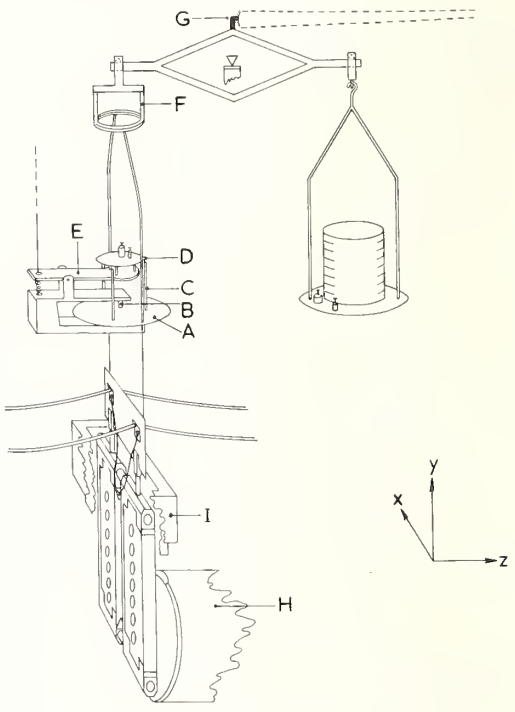


FIGURE 2

vertical wires are joined by upper and lower flexible ribbons of wires B, and are terminated on lugs on the plate K from which the coil is suspended. The coil so formed had 32 turns and was 25 mm broad and 500 mm long, and had a resistance of about 3.7 Ω.

The girders were kept at one of three distances apart equal to about 5, 29 and 53 mm by one of three pairs of fused silica spacers D of equal length which have three raised bosses polished flat and parallel to 1/10 fringe at each end. These are kept in contact with silica plates C cemented into the girders by springs E. The lengths of the spacers were measured by wringing etalon plates to either end and using the method of exact fractions on fringes formed in light of standard wavelengths, including the ^{86}Kr standard.

The coil could be adjusted to be vertical by the mechanism M. The coil assembly was very rigid. It weighed 1.7 kg, and was suspended by the wires L from the pan of a balance as shown in figure 2.

The force change on reversal of the current flowing in the coil was opposed by lowering or raising calibrated stainless steel weights on a small scalepan D by the mechanism E. The balance is not arrested during this operation, and the knife-edges remain in contact with their planes under constant load. Excessive travel is prevented by the pan-stops B. The pan is supported from its knife-edge stirrup by flexible phosphor bronze strips F, and the balance rest point is read by the mirror G and a lamp and scale. The mass of the coil was counterpoised by a mass of similar density to prevent changes in barometric pressure causing the rest point to drift.

The balance is mounted on a subplate which can be moved smoothly enough, even when the beam is not arrested, to adjust the position of the coil in the magnet gap. This plate rested on a cantilevered platform fastened to a concrete pillar which went deeply into the subsoil. The balance case was lagged, and the coil protected by cardboard draught screens, and the whole apparatus was inside a thermally lagged cabin which was itself inside a temperature-controlled room. All operations connected with determining the force on the coil were carried out from outside this cabin.

The coil hung with its lower third between the polefaces of a magnet H which were adjusted to lie in a vertical plane by a set-square and spirit level. Mu-metal shielding I surrounding the upper portion of the coil reduced the effect of stray magnetic fields acting on the top of the coil. The damping action of the girder metal interacting with the field of the magnet altered the 25 second period free swing of the balance to a 55 second time constant damped logarithmic motion.

Rose shims fitted to the pole-faces had plane surfaces inclined at 30° to the horizontal machined from the top to reduce x-components of the field at the gap edge. These interact with the lack of straightness in the y–z plane of the vertical currents to produce unwanted forces.

The magnet was energised by a regulated current source and a field stability of 1 ppm/10 min was achieved. The nonuniformity of the field at the centre of the gap was mapped from time to time during the experiment with the NMR probe used to determine the precessional frequency, and a typical

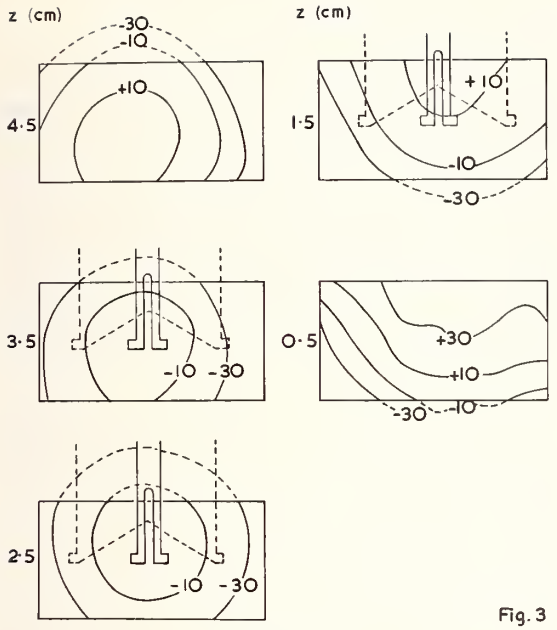


FIGURE 3

example is shown in figure 3. The contours are 20 ppm apart, drawn in five planes at 1 cm intervals across the magnet gap, and the full and dashed lines represent the courses of the lower coil conductors with respectively the shortest and longest spacers in position.

The precessional frequency of the protons in a sealed cylindrical sample of water doped with $FeCl_3$ having less than 2×10^{19} ions/cm^3 was measured by the simple arrangement shown in figure 4. Absorption of energy on resonance with the 20 MHz oscillations applied to the tuned circuit was detected and amplified to provide the vertical deflection of an oscilloscope. The horizontal deflection was provided by the same oscillator that modulated the magnetic field at 10 Hz to sweep the resonant condition, via a passive phase-shift network to separate the forward and return traces and provide a symmetrical crossover defining the centre of the absorption.

The lineshape varied with the position of the probe in the magnet gap, being symmetrical near the centre,

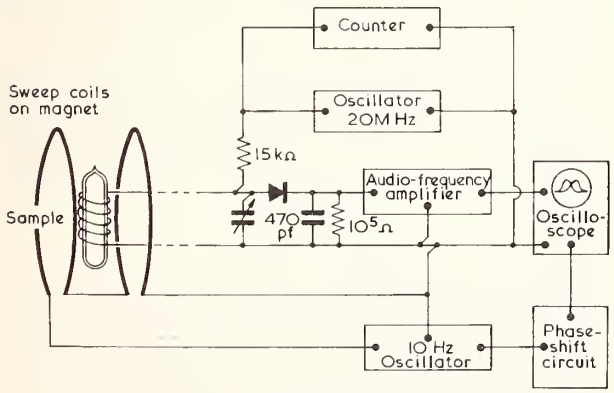


FIGURE 4

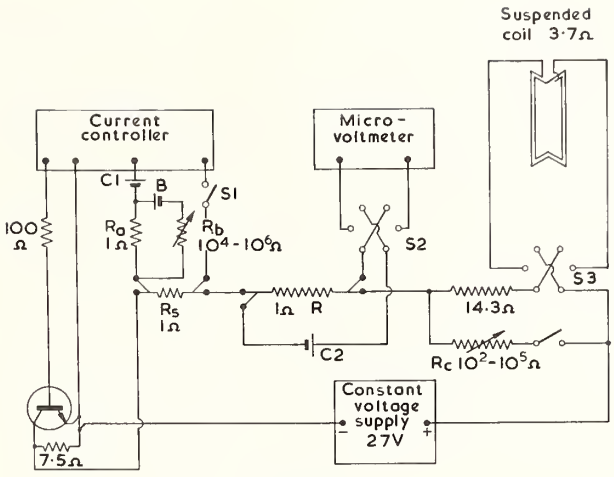


FIGURE 5

and so all field values are referred to this point, particularly the displacement of the unsymmetrical signal when the probe was in the position it occupied during weighings. The inhomogeneous field from the current in the suspended coil prevented observation of the signal, and so it was observed whilst this current was zero, during the reversal procedure.

The current was regulated and measured to 1 ppm by the circuit shown in figure 5, R being a well-aged stable resistor, and C2 a standard cell which was compared daily with a battery of twelve cells. The cells and resistor were in turn compared with the maintained NPL volt and ohm before and after the experiment. The resistor R_c shunted a known fraction of the current through the suspended coil to bring the balance rest point close to an arbitrary zero. This rendered the use of riders or small weights unnecessary and minimised error caused by the determined balance sensitivity.

3. Experimental Procedure

The balance rest-point was obtained with the current flowing and again with the current reversed, the weights being raised or lowered as necessary; again with the current in the reversed direction and again in the original direction. This procedure eliminates the effect of drift in the experimental conditions. During three intervals whilst the current was stopped as well as at the start and end the precessional frequency was observed, and whilst the rest points were being obtained the current was measured and the orientation of the coil about the y and z axes observed by a telescope which views scales reflected in one of the silica plates set in the girders.

The whole procedure was repeated with the balance rotated to provide different orientations of the coil about the y axis, so that the direction and magnitude of the maximum force on the coil could be deduced and corrected back to the direction in which the spacers are perpendicular to the field, since this is the direction for which the change in width of the

coil is defined. At no time was the balance arrested during this weighing.

The balance sensitivity was determined in both current directions by shunting the coil with known values of R_c to produce force changes deducible from the approximate value of the total force, and observing the change in the rest point.

Various auxiliary information including the barometric pressure and temperature in the vicinity of the calibrated masses (in order to make a correction for air buoyancy) was gathered, and then the coil was dismantled and reassembled with a different spacing, and another weighing performed.

Five sets of weighings in all were carried out during the period 16th June–24th July 1970. Each set consisted of three weighings with each spacer. For the first set the experimental technique was being perfected, and the second set was a repeat with higher precision. The coil was rotated 180° about the vertical axis for the third set, inverted for the fourth set, and again rotated whilst inverted for the fifth set. Inversion, and to a lesser extent, rotation, alters the interaction of the measured values of B_x and B_z along the routes of the vertical wires with the measured lack of straightness of these. The difference between the results for these sets, which agrees with our estimates, reflects the magnitude of this interaction.

Less than 3×10^{-4} N of the force on the coil resulted from the interaction of the coil with the magnetism it itself induces in the soft iron of the electromagnet and in the mu-metal shielding, but this force does not reverse direction with reversal of the current, and so it does not affect our result. It was responsible for some part of one unwelcome effect, however, in that the balance sensitivity could change by as much as 50 percent with the direction of the current in the coil. This change could be made very small by precise positioning of the coil in the z direction within the magnet gap, and this was done for all the results reported here. A force change indistinguishable from that obtained with the coil so positioned was found for a single weighing with the coil positioned to produce a large sensitivity change. This weighing was carried out with the largest spacing for which the difference in sensitivity was most pronounced. We take this to mean that there would not be a significant effect on our results from the much smaller random differences in sensitivity obtained by precise coil positioning.

4. The Results

A least squares fit to the best straight line of coil spacing versus force obtained for that spacing was calculated for each of the four final sets of results. From the slope and the precession frequency to which all the results were corrected, a value of γ_p' was deduced. The departures of each set from the mean result were -27, $+14$, $+14$, and 0 ppm, whereas one standard deviation of each set obtained by the least squares analysis on the straight line was 9, 7, 12, and 17 respectively.

The final result in terms of the NPL maintained ampere of June–July 1970 is

$$2.675075(43)\times10^8 \text{ rad s}^{-1}\ T^{-1} \qquad (16 \text{ ppm})$$

and expressed in terms of the absolute ampere as measured in 1970 [5] is

$$2.675082(45)\times10^8 \text{ rad s}^{-1}\ T^{-1} \qquad (17 \text{ ppm})$$

which may be compared with the recommended value quoted by Taylor, Parker and Langenberg [6]

$$\gamma_p' = 2.6751270(82)\times10^8 \text{ rad s}^{-1}\ T^{-1} \qquad (3.1 \text{ ppm}).$$

The figures in parentheses are one standard deviation in the last two digits of the quoted value.

5. Conclusion

It must be emphasized again that the results reported above were obtained with an apparatus assembled only for the purpose of gaining experience with the technique. In particular, the poor homogeneity of the high magnetic field was not very reproducible, the stray field at the top of the coil was variable in magnitude and direction, the vertical sides of the coil were not sufficiently straight and the magnetic field components perpendicular to them through which they passed were too large, and we did not carry out our avowed intention of measuring the coil spacing interferometrically *in situ*. Nevertheless we felt the result obtained was of sufficient accuracy to be worth reporting at this stage. Furthermore it has encouraged us in the belief that the method is sound, and that when the above four difficulties are overcome, a result to a much higher accuracy approaching 1 ppm can be attained.

We feel also that we have demonstrated the advantages of working with a coil of variable separation. The most important of these are:

(1) the length determination is carried out directly in terms of the ^{86}Kr length standard; in an improved apparatus this would take place simultaneously with the determination of the force on the coil,

(2) elimination of systematic errors arising from the unknown current distribution in the conductors,

(3) elimination of the need to measure tediously the position of every point to the highest precision on metrologically difficult objects, the vertical wires,

(4) elimination of the effect of slight distortion of the coil former caused by the heating effect of a current flowing in the coil, since this is the same for each coil spacing.

A more complete description of this experiment is available as an internal report of the National Physical Laboratory, of whose research programme this determination is a part.

References

[1] Vigoureux, P., Metrologia **1**, 3 (1965).
[2] Vigoureux, P., Proc. Roy. Soc. (A) **270**, 72 (1962).
[3] Captuller, H., Zeit. Instrumentenkunde **69**, 191 (1961).
[4] Captuller, H., Zeit. Instrumentenkunde **69**, 133 (1961).
[5] Vigoureux, P., private communication.
[6] Taylor, B. N., Parker, W. H., and Langenberg, D. N., The Fundamental Constants and Quantum Electrodynamics, (Academic Press, New York, 1969).

DISCUSSION

H. Capptuller: We thank Dr. Kibble for his very interesting paper on this matter, and we see the differences between the two methods. One of these methods or the other has some difficulties, and the future will show what method will be the best and what method will lead to the best results.

H. G. Robinson: I was wondering if you had estimated any of the effects due to the Araldite on the measurements. Are there any systematic effects due to that?

B. P. Kibble: You mean that it's magnetically permeable?

H. G. Robinson: Yes. Changing the magnetic field gradient.

B. P. Kibble: I have to admit no. We used as little as possible if that's any help. I would think less than a thousandth of an inch thick layer. But do you have any figures or experience that show this is bad? The coil, when carrying no current, was put near the NMR probe and the NMR signal didn't shift by the smallest extent that we could observe, about half a part per million, and I think this is a reasonable experimental check for all the materials in the coil. The girder material was checked separately. It has a susceptibility of 0.0000016 which might change the field in the vicinity by two to three ppm, and in an experiment of higher precision than that we'd have to, of course, again measure the actual change in field experimentally by mapping or use a different material for the girders.

FARADAY CONSTANT AND ATOMIC MASSES

The Faraday and Its Significance in Determining the Fundamental Constants

A. Horsfield

National Physical Laboratory, Teddington, Middlesex, United Kingdom

The principles of the electrolysis of solutions of electrolytes are briefly discussed and the Faraday is defined. Experimental methods are given, particularly details of precise determinations of the electrochemical equivalents of silver and of iodine, used in the determination of the Faraday. The accepted value of the Faraday (1970) is 96,487.0 ± 1.6 coulombs gram-equivalent^{-1}. The value of the Faraday, which is the product of the charge of the electron and Avogadro's number, is important in the evaluation of the fundamental physical constants and its role is summarized. Suggestions are made for future experimental work.

Key words: Electrochemical equivalent; electrolysis; Faraday; fundamental constants.

1. Early Electrochemistry

Study of the conduction of electricity in liquids became possible at the beginning of the nineteenth century, following the discovery of the electrolytic cell by Volta in 1800, which provided the first continuous source of electric current. It was soon discovered that the conduction of electricity by solutions is accompanied by chemical reactions at the electrodes which serve to conduct the current into and out of the solution. Nicholson and Carlisle demonstrated the decomposition of water into hydrogen and oxygen by a current in 1801. Davy's discovery of sodium and potassium metals by electrolysis of

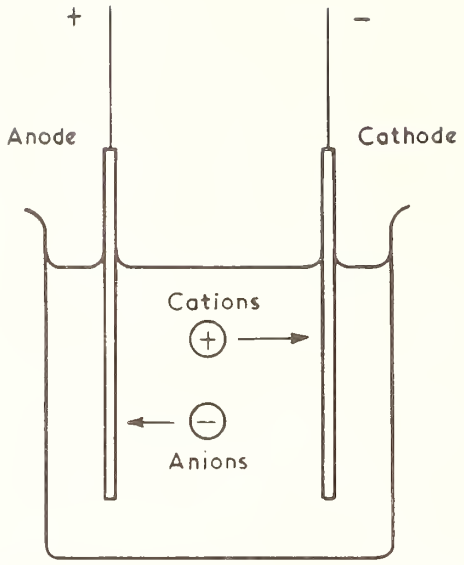


FIGURE 1. *Diagram illustrating the working of the basic coulometer.*

moist soda and potash was a striking example of the novelty of electrochemical decomposition. Many of the phenomena of electrolysis were already known when Michael Faraday began his researches.

Figure 1 illustrates a simple electrolytic apparatus or 'coulometer' (often called 'voltameter' in older literature). The positive and negative electrodes are respectively designated anode and cathode. The electrolyte, usually a simple metal salt, molten or in solution, dissociates into charged ions which carry the current through the solution. The positive cations move to the cathode under the influence of the potential gradient between the electrodes while the negative anions move to the anode. The ionic species move independently and the fraction of the total current carried by each depends on the charge of the ion involved and on its velocity. The nomenclature was introduced by Faraday.

It was the quantitative relationship between electrochemical change and current which interested Faraday and enabled him to correlate the mass of experimental data that had accumulated since 1800. Faraday's laws of electrolysis, which were published in 1833, state:

(1) that the amount of chemical decomposition produced by an electric current (that is, the mass of substance deposited or dissolved at an electrode) is proportional to the quantity of electricity passed.

(2) that the amounts of different substances released or dissolved at electrodes by the same quantity of electricity are proportional to their chemical equivalents.

From the second law, it follows that the amount of electricity required to liberate or dissolve one equivalent weight of any substance by electrolysis is constant.

It was not until after Faraday's death that the significance of his laws of electrolysis for atomic theory was realized. In 1881 von Helmholtz pointed out that if elementary substances are composed of atoms, it follows from Faraday's laws of electrolysis that electricity also is composed of elementary portions which behave like atoms of electricity. Investigations on the conduction of electricity by gases led to the identification of the electron as the fundamental unit of electricity at the end of the century. Faraday's positive and negative ions are therefore atoms (or groups of atoms or radicals) with a deficiency or an excess of an integral number of electrons, where the integral number is the valency of the atom. The ions move in opposite directions through the solution to the electrodes where their charges are neutralised, causing them to be discharged to neutral atoms or radicals.[1] These are the primary electrode reactions, of which the deposition of silver on a platinum cathode in the silver coulometer is a typical example. The discharged atoms may, however, react with the solvent or with the material of the electrode to give secondary reaction products. These products will also appear in quantities proportional to their chemical equivalents, provided the secondary reactions go to completion.

2. The Faraday

Quantity of electricity or charge is measured in coulombs, the product of the current in amperes and the time in seconds for which it flows. If the electrochemical equivalent of a substance, E, is the mass of substance deposited or dissolved by the passage of one coulomb during electrolysis, then from Faraday's laws the total mass of substance, W, involved in electrochemical reaction by a current, I, acting for a time, t, is:

$$W = EIt. \tag{1}$$

The quantity of electricity required to react one chemical equivalent of a substance is called the Faraday, F, and it is a universal constant. Thus eq (1) becomes:

$$A/z = EF \tag{2}$$

where the chemical equivalent of the substance is given by the ratio of its atomic weight, A, to its valency, z.

From eqs (1) and (2), the measurement of the Faraday involves the determination of the electrochemical equivalent from the mass of substance deposited by an arbitrary charge, and the determination of the atomic weight. The precise determination of atomic weights, from nuclidic masses and their abundances, will not be considered here and the reader is referred to the review by Professor A. H. Wapstra, these Proceedings.

[1] Charged radicals are not normally neutralized directly at the electrodes; instead, other processes occur which involve the solvent or hydrogen ions because they are favoured thermodynamically. The details of the electrode reaction mechanisms are not directly relevant to a discussion of the Faraday and, because the net chemical changes at the electrodes may be represented by the neutralization of the charged radicals, this approach is adopted for simplicity in this paper.

3. Measurement of Electrochemical Equivalents

3.1. The Coulometer

The traditional coulometer employed deposition of the discharged metallic ions on the cathode from a solution of an electrolyte, MX. The cathode was either the metal, M, itself or an inert noble metal such as platinum, Anodes of the same material, M, were preferred since the electrode reaction involving the anions resulted in regeneration of the electrolyte, MX, by erosion of the metal; the use of platinum anodes with aqueous electrolytes causes oxygen to be evolved and the solution becomes depleted of metal cations, M^+, as the electrolysis proceeds. Denoting the electron by '*e*', the chemical changes occurring at the electrodes may be represented as:

at the cathode $M^+ + e \rightarrow M$

at the anode $X^- + M \rightarrow MX + e.$

The amount of electrochemical action was obtained from the increase in the mass of the cathode due to the deposit which, by Faraday's first law, is equal to the decrease in mass of the anode.

The simple form of coulometer in figure 1 is unsuitable for precise work since small particles tend to fall from the electrodes during electrolysis, particularly from the anode (anode sediment). Early workers in the field overcame this disadvantage by adopting platinum bowls for the cathodes to contain all the deposited material and by effectively separating the electrodes in two compartments to prevent anode sediment from contaminating the cathode deposit, thereby giving a false value for the increase in mass. Figure 2 shows some of the early forms of silver coulometer and illustrates the use of filter paper (Rayleigh and Sidgwick [1]), porous pots

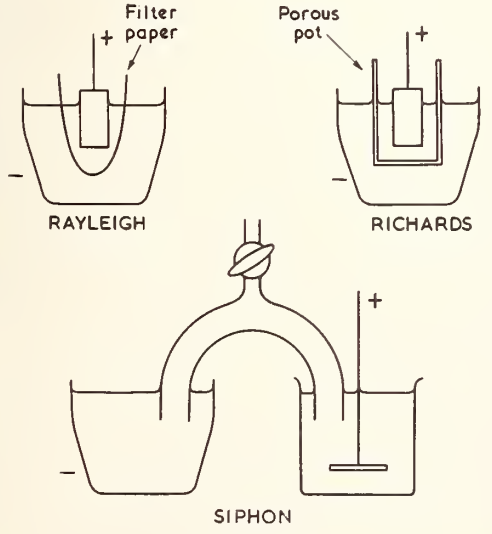


FIGURE 2. *Early forms of silver coulometer, showing methods used to prevent anode sediment from contaminating the deposit on the platinum bowls used as cathodes.*

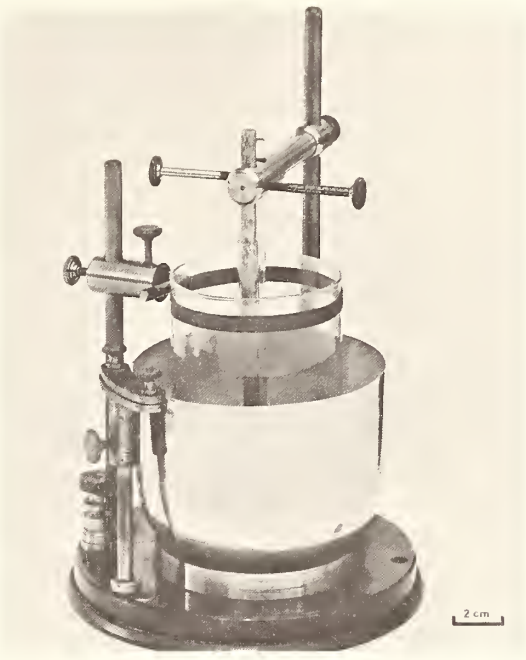


FIGURE 3. *The NPL silver coulometer (1910), designed by F. E. Smith, which was used to realize the international ampere.*

(Richards and Heimrod [2]) and siphons to produce separate electrode compartments. Other coulometers are illustrated in figures 3, 5, and 6.

3.2. The Coulometer Circuit

A typical electrical circuit used in coulometry is drawn in figure 4. The current is adjusted by the variable resistance, V, until the potential difference across a standard resistance, R, just balances the emf of a standard Weston cell, E, so that no current is observed to flow through the galvanometer, G, when the switch S_3 is closed. Since the values of R and E can be checked to about 1 part in 10^6 (1 ppm), the value of the current can be obtained from the relation:

$$I = E/R. \qquad (5)$$

The potential difference across the standard resistance can also be checked using a potentiometer which has been calibrated against a second standard cell. In practice, it is necessary to minimise dissipation of heat by the current in the solution, to prevent large temperature rises, and currents in the range $\frac{1}{2}$ to 1/20 ampere have been used in precise work.

D is a dummy load, equal to the resistance of the coulometer. The switch S_1 enables the current to be diverted from the dummy to the coulometer for a timed run. A timing circuit, calibrated by a standard clock, is connected to the switching device, S_2, which is operated by the current passing through the coulometer so that the exact duration of the electrolysis can be measured. A run of two hours can be timed with an uncertainty of less than 1 ppm.

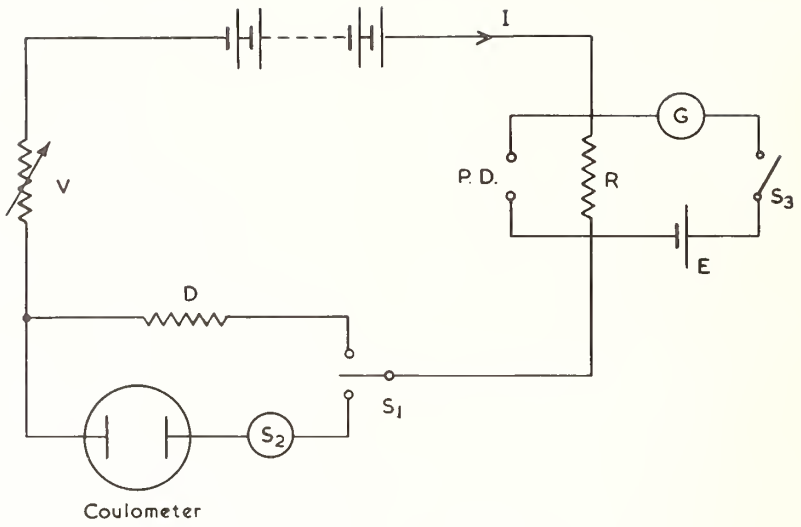


FIGURE 4. *The basic electric circuit used in coulometry.*

The current obtained from eq (5) is the value of the current as 'maintained' by standard cells and resistances. This can only be related to the ampere,[2] which is defined in terms of the forces between conductors carrying currents, by the use of a current balance which gives the ratio, K, of the 'maintained' to the 'absolute' ampere. These comparisons are performed at the national standardising laboratories. In 1965, Vigoureux [6] determined the value of K_{NPL} to be 1.0000157 ± 3.6 ppm for the ampere as maintained at the National Physical Laboratory (NPL). For precise work therefore, eq (5) should be modified to:

$$I_{abs} = E/KR. \qquad (6)$$

4. Early Measurements

4.1. The Silver Coulometer

One of the first systematic studies with the silver coulometer was by Lord Rayleigh and Mrs. Sidgwick in 1884 [1]. They used the form shown in figure 2, with a platinum bowl as the cathode, sheet silver wrapped in filter paper as the anode and aqueous silver nitrate for the electrolyte. They measured the current with a current balance and obtained 1.11794 milligram coulomb^{-1} (mgC^{-1}) for the electrochemical equivalent of silver. Among other precautions, they heated the platinum bowl with its deposit before the final weighing, in an attempt to eliminate possible 'inclusions' of solvent or silver nitrate that might be formed in the silver deposit during electrolysis. Rayleigh and Sidgwick varied the conditions of electrolysis and found, for example, excess deposits when silver acetate was added to the electrolyte and heavier deposits (by about 0.02%) in large bowls than in their small ones. These effects were only satisfactorily explained by later investigators [8, 9].

F. and W. Kohlrausch [3] published a determination of the electrochemical equivalent of silver in 1884. The current in their coulometers was determined with tangent galvanometers and the value obtained was 1.1183 mg C^{-1}.

4.2. Deficiencies of the Silver Coulometer

Faraday was among the first to realize that electrochemical decomposition offered a convenient way to measure electric current. Following the detailed studies of the silver coulometer by Rayleigh and Sidgwick and by the Kohlrauschs, attention was almost entirely confined to this form. The silver coulometer was chosen as the official primary standard for the measurement of current by the International Electrical Congress at Chicago in 1893 and the choice was confirmed at the International Electrical Conference of 1908 in London. The 'international' ampere was defined as the steady current causing silver to be deposited at the cathode of a silver coulometer at the rate of 1.11800 mg S^{-1}, thereby fixing the electrochemical equivalent of silver. Interest in coulometry therefore shifted from precise determinations of electrochemical equivalents and the Faraday, to establishing forms of coulometers and conditions of electrolysis which gave consistent and reproducible results. With the defined value of the electrochemical equivalent of silver, the coulometer was used to establish the ampere so that the experiments described in section 3 could be used in reverse to calibrate the emf of the standard Weston cell which, with the standard resistance, was used to maintain the ampere on a day-to-day basis.

In the years before the World War I, considerable efforts were made to resolve the discrepancies in the masses of silver deposited in different forms of coulometer by the same current, which could amount to more than 0.01 percent in extreme cases. The various

[2] This was formerly designated the 'absolute' ampere in contrast to the 'international' ampere (defined in terms of the silver coulometer). The international electrical units were abolished in 1948 by international agreement.

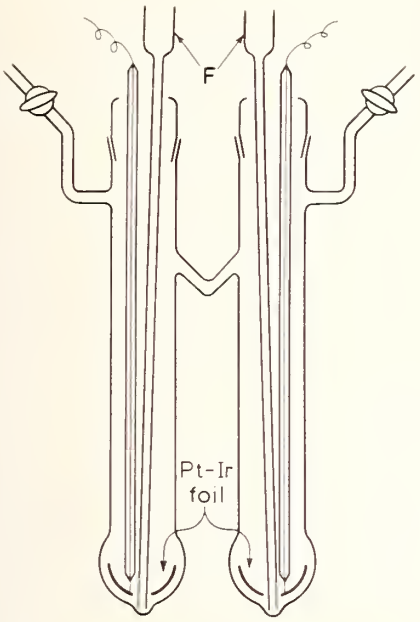


FIGURE 5. *The iodine coulometer used by Washburn and Bates (1912).*

national standards laboratories were active in this field, for example Guthe [4] and Rosa, Vinal and McDaniel [8, 9, 10, 11] at the National Bureau of Standards (NBS) and Smith, Mather, and Lowry [5] at the NPL.

It was found that silver deposits from coulometers containing filter paper were heavier than from those without. They were also observed to be striated, imperfectly crystalline and of tarnished colour, whereas the deposits from coulometers in which porous pots were used to separate the electrodes had the lowest masses and consisted of well-formed crystals of white colour. Detailed researches by Rosa, Vinal, and McDaniel [8, 9] showed that filter paper is not inert but produces organic impurities in the solution of electrolyte that are believed to give rise to colloidal silver; this is deposited at the cathode, increasing the weight of silver and breaking up the regular growth of the deposited crystals. The Rayleigh form of coulometer was discarded for accurate work, and the use of filter paper in preparing the electrolyte was also avoided.

Another major problem in accurate work with the silver coulometer was the discovery that inclusions of solvent and silver nitrate could occur in the deposited silver. Rayleigh and Sidgwick had looked for this effect inconclusively but later workers found that inclusions could increase the weight of the silver deposit from 0.0008 percent to 0.030 percent, depending on the form of the coulometer and conditions of electrolysis. The principal method used in estimating inclusions were to find the decrease in weight after heating the platinum basin with its deposit to red heat. Attempts were also made to measure the precise mass of silver in the weighed deposit by chemical analysis. Although it was concluded that the amount of inclusions in the silver, under conditions of electrolysis which yield a highly crystalline deposit, does not exceed 0.005 percent by weight [11], the possibility of inclusions remains a serious source of error in all coulometric work in which solid deposits are used. The subject of inclusions has been reviewed by Scott [12].

The purity of the materials used is vitally important, both of the electrolyte and also of the electrodes from which impurities can be brought into solution. Ionic impurities in the solution carry a fraction of the current but do not make a proper contribution to the mass of the deposit which is assumed to be entirely silver. Besides the effects of organic impurities already noted, it was found that minute quantities of reducing impurities, colloidal silver and small amounts (parts per million) of free acid or base in the electrolyte could affect the weight of the deposit by as much as 0.009 percent [11]. It was found possible to obtain reproducible deposits by using silver nitrate purified by recrystallisation and fusion and electroformed silver anodes. The various forms of coulometer (excluding the Rayleigh form) agreed with one another to within 20 ppm when all precautions were observed [11] and this was considered satisfactory, in 1910, for a primary standard of current. A comprehensive review of the early work on the silver coulometer has been given by Rosa, Vinal, and McDaniel [8, 9, 10, 11].

4.3. The Iodine Coulometer

While work was in progress to perfect the silver coulometer as a means of establishing the international ampere, Washburn and Bates [13] made a study of the iodine coulometer at the University of Illinois.

The iodine coulometer is based on the chemical reaction:

$$I_3^- + 2e \rightleftharpoons 3I^- \qquad (7)$$

which proceeds in the forward direction (left to right) at the cathode and in the reverse direction at the anode. The tri-iodide ion, I_3^-, is formed by combination of iodine with the mono-iodide ion:

$$I_2 + I^- \rightarrow I_3^- \qquad (8)$$

and the molecular iodine, which would otherwise be formed at the anode, remains in solution as the soluble tri-iodide ion.

Figure 5 shows a diagram of the iodine coulometer. The electrodes were platinum-iridium alloy which is not attacked by iodine. The anode and cathode compartments, filled with 10 percent potassium iodide solution, were separated by a V-shaped connection. About 25 ml of concentrated potassium iodide solution was admitted through the capillary funnel, F, to cover the anode; it was added to provide a sufficient source of iodide ion to be oxidised to iodine. A weighed amount of a standardised solution of iodine in potassium iodide was introduced at the cathode in the same way to provide a source of iodine to be reduced to iodide ions.

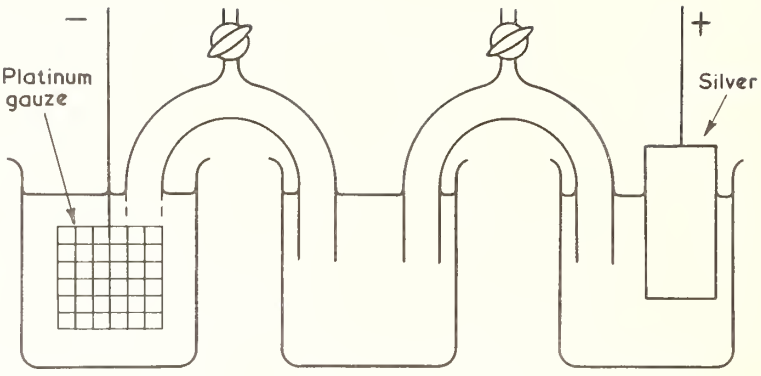


FIGURE 6. *The form of silver coulometer used by Craig, Hoffman, Law and Hamer (1960) in the anode dissolution determination of the Faraday.*

On electrolysis, iodine was consumed at the cathode and liberated at the anode, according to eq (7), resulting in a net transfer of iodine from the cathode to the anode. At the end of the run, the solution round each eleetrode was siphoned off through the capillary funnels, F, and the amount of iodine at each electrode was determined by analysis with solutions of pure arsenious acid whieh has previously been standardised by titration against weighed quantities of pure iodine. The analysis is based on the reaction:

$$H_3AsO_3+I_3^-+H_2O\rightarrow H_3AsO_4+2H^++3I^- \quad (9)$$

which occurs between iodine and arsenious acid.

By comparing the loss of iodine at the cathode and the gain at the anode, which should be equal, it was shown that reaction (7) is reproducible and reversible within the experimental error of about 20 ppm.

This eoulometer meets the principal objections to the silver coulometer:

(1) there is neither anode sediment nor loose particles from the cathode
(2) there are no inclusions to be estimated since the iodine released at the anode remains in solution.

The only disadvantage of the iodine method is that precise, quantitative chemical analysis is tedious and far less convenient than finding the change in mass of an electrode by weighing.

The electrochemical equivalent of iodine was found from a direct comparison with silver by Vinal and Bates [14] at the NBS. The same current was passed through the iodine coulometer in series with a silver coulometer and the ratio of the masses of iodine released to silver deposited gave the ratio of their electrochemical equivalents directly without reference to the current passed through the coulometers. The value 1.11800 mg C^{-1} was assumed for the electrochemical equivalent of silver and the value 1.31502 mg C^{-1} was obtained for iodine.

5. Recent Measurements of the Faraday

No precise work on the determination of electrochemical equivalents or the Faraday appears to have been done after 1914, when the iodine work was published, for nearly forty years. The silver coulometer, under specified conditions, continued to be used as the primary standard of international current and from time to time it was employed as a means of calibrating the emfs of standard cells used to maintain the international volt, for example, as in the international comparison at Berlin in 1932 [15]. It was not until 1953 that any new work was done on the Faraday, when the oxalate coulometer was devised at NBS by Craig and Hoffman [16].

5.1. The Oxalate Coulometer

The coulometer used for the oxalate determination was similar to that used in the later silver dissolution work at the NBS and shown in figure 6. In the oxalate experiment the cathode was platinum gauze but the anode was made from perforated gold sheet and the electrolyte was sodium oxalate of NBS-certified purity in a supporting electrolyte of dilute sulphuric acid. The anode reaction was:

$$(CO_2)_2^{--}\rightarrow 2CO_2+2e \quad (10)$$

involving a loss of oxalate anion by decomposition to carbon dioxide.

A weighed quantity of sodium oxalate was added to the anode compartment and a known current was passed for a measured time. The mass of oxalate remaining after electrolysis was determined by chemical analysis which involved titrating with standardised potassium permanganate solution against the anode solution; the amount of oxalate consumed in electrolysis was obtained by difference. The coulombs passed were measured in the way already described in section 3.2.

The electrochemical equivalent of the oxalate anion and hence the value of the Faraday were obtained. This method, like the iodine work, was free from the disadvantages of anode sediment and inclusions which are inherent to the silver deposition coulometer.

5.2. The Silver Dissolution Method

A further redetermination of the Faraday at the NBS, using a silver coulometer once more, was

published in 1960 by Craig, Hoffman, Law, and Hamer [17]. This is the most recent work and it is a model of thorough, metrological experiment.

To overcome the disadvantages of the classical silver coulometer—the formation of inclusions in the deposit on the cathode and the possibility of partial separation of the isotopes ^{107}Ag and ^{109}Ag during electrolysis—the procedure of measuring the amount of erosion of the anode was adopted since electrolytic dissolution of the anode is independent of both effects.

The apparatus used is shown in figure 6. It is a modified siphon form of coulometer to ensure adequate separation of the electrode compartments. The electrolyte used was 20 percent perchloric acid containing $\frac{1}{2}$ percent of silver perchlorate to reduce the spontaneous solubility of the anode silver in the solution to a negligible value. The anodes were pure silver, demonstrated by spectrochemical analysis to have a total of less than 1 ppm of impurities. They were treated to remove oxygen and allowed to crystallise from the melt under vacuum. It is probable that all the anodes used were of single crystal form [17]. The cathode was platinum gauze. The electrical circuit was essentially that shown in figure 4.

Numerous variations in conditions—current density, anode potential and anodes with different histories of purification procedure—were made in a search for systematic errors. Anode sediment was carefully collected and weighed and corrections were applied to the mass of silver dissolved. Corrections, although small, were also made for the known impurities in the silver.

The mean electrochemical equivalent of silver, from 31 determinations was 1.117972 ± 0.000019 mg C^{-1}. The atomic weight of silver was redetermined at the NBS by mass spectrometry, during which work it was also established that no measurable separation of the silver isotopes occurs during electrolysis. Using the value 107.9028 ± 0.0013 found for the atomic weight of silver on the physical scale, the value obtained for the Faraday—'the new silver value'—was $96{,}516.5 \pm 2.4$ coulomb gram-equivalent^{-1} (Cg-equiv.$^{-1}$) on the physical scale.

6. Value of the Faraday

In table 1 are listed some determinations of the electrochemical equivalent of silver, in which the current was measured by some form of current balance, giving a result independent of the emfs of standard cells which were subject to great uncertainty at the turn of the century. The unweighted average of these values is 1.11806 mg C^{-1} which, if the correction for inclusions [11] of -0.004 percent be applied, becomes 1.11802. The value recommended by the International Electrical Congresses of 1893 and 1908 was 1.11800 mg C^{-1}, and this differs from the new silver value of Craig et al. [17] by only 30 ppm. When it is considered that in 1908 the absolute ampere could be realized only with a probable error of 3 parts in 10^5 [18] the choice of 1.11800 mg C^{-1} for the defined value of the electrochemical equivalent for silver must be considered as being in remarkable agreement with modern work. The value obtained for the Faraday, taking the accepted atomic weight of silver at that time of 107.88 was 96,494 C_{int} (g-equiv)$^{-1}$.

Table 1. *Values obtained for the electrochemical equivalent of silver*

Electrochemical Equivalent (mg C^{-1})	Reference	Date
1.11794	Rayleigh and Sidgwick [1]	1884
1.1183	Kohlrausch and Kohlrausch [3]	1884
1.11773	Guthe [4]	1904
1.11827 ± 0.00002	Smith, Mather and Lowry [5]	1908
1.117972 ± 0.000019	Craig, Hoffman, Law and Hamer [17]	1960
1.11800 (defined)	Chicago Congress	1893
	London Conference	1908
1.11805 ± 0.00007 (adjusted)	Birge [19]	1929

From the value of the electrochemical equivalent of iodine, obtained from their comparison between the silver and iodine coulometers, Vinal and Bates [14] found 96,515 C_{int} g-equiv^{-1} for the Faraday. The difference between this and the silver value of 96,494 C_{int} g-equiv^{-1}, amounting to 220 ppm, could not be satisfactorily accounted for and a rounded average of $96{,}500 \pm 10$ C g-equiv^{-1} was therefore recommended by Vinal and Bates for general use.

In his first evaluation of the fundamental constants, Birge [19] found that the value for the atomic weight of iodine implied by acceptance of the values for the electrochemical equivalent of iodine and the Faraday given by Vinal and Bates did not agree with the best value for the atomic weight of iodine available in 1929. He therefore retained the silver values which, converted to absolute units, became 1.11805 mg C^{-1} for the electrochemical equivalent of silver and $96{,}489 \pm 7$ C g-equiv^{-1} for the Faraday.

The value of the iodine Faraday, as revised by Vinal and Hipple [20] received support from the work of Sommer and Hipple [21] who measured the constant by an independent, physical method in-

Table 2. *Values obtained for the Faraday, and revised by Hamer [23] using the unified ^{12}C scale of atomic weights and absolute electrical units*

Date	Method	Value (C g-equiv^{-1})	Revised value (1968)
1908	Silver deposition	96 494	96 478
1914	Iodine oxidation	96 515	$96\,489 \pm 3$
1953	Oxalate oxidation	$96\,518 \pm 3$	$96\,482.8 \pm 3$
1953	Electromagnetic	$96\,522 \pm 3$	$96\,486.1 \pm 0.9$
1960	Silver dissolution	$96\,516.5 \pm 2.4$	$96\,486.6 \pm 1.6$
1969	Adjusted value [7]		$96\,486.7 \pm 0.54$

volving cyclotron resonance of protons in 1953. The magnetic moment of the proton, in nuclear magnetons, μ_p/μ_n, is given by the ratio of the cyclotron resonance frequency to the magnetic resonance frequency of protons in the same magnetic field. The gyromagnetic ratio of the proton, γ_p, is obtained from the precession frequency of protons in a known magnetic field. The Faraday is related to both atomic constants by the equation

$$F=\gamma_p M_p(\mu_p/\mu_p)^{-1} \tag{11}$$

where M_p is the mass of the proton. Using their value of μ_p/μ_n and an earlier value for γ_p, Sommer and Hipple obtained 96,522±3 C g-equiv^{-1} in excellent agreement with the revised iodine Faraday of 96,521.5 C g-equiv^{-1}.

Although much meticulous and painstaking work has been devoted to the evaluation of the Faraday over the years, the new value of Craig, Hoffman, Law, and Hamer [17] by the silver dissolution method must be given precedence over the older work because of modern advances in the preparation and analysis of materials. Their silver was demonstrated to have a total impurity content of less than 1 ppm. The new silver value has therefore been adopted exclusively in evaluations of the fundamental physical constants that have appeared since 1960.

The discrepancy between the silver and iodine Faradays remained unresolved until 1968 when original data of Vinal and Bates was reexamined by Zielen [22]. He considered only those experiments in which the total charge had been measured in addition to the iodine released at the anode of the iodine coulometer and the silver deposited in the silver coulometer. Ignoring the silver deposit, Zielen reworked the iodine data according to eqs (1) and (2), using the 1961 value for the atomic weight of iodine, and found agreement with the new silver value of the Faraday to 3 ppm. Although this result is very satisfactory, it must be to some extent regarded as fortuitous since the realization of the electrical units was subject to an error of some 30 ppm in 1914. In a similar way he also satisfactorily resolved the disagreement between the new silver Faraday and the electromagnetic result of Sommer and Hipple. These conclusions were confirmed by Hamer [23] in an independent reevaluation of the older work.

The value of the Faraday recommended by several authors [7, 22, 23] is 96,487.0±1.6 C g-equiv^{-1} and the adjusted magnitude of the Faraday resulting from the latest least-squares adjustment of the fundamental physical constants by Taylor, Parker, and Langenberg [7] of 96,486.7±0.54 C g-equiv^{-1} is in agreement.

7. Significance of the Faraday for the Fundamental Constants

The Faraday represents the charge carried by one chemical equivalent of ions; for singly charged ions the chemical equivalent is identical with the atomic weight and hence is the Faraday is the product of Avogadro's number and the charge of the electron.

$$F=Ne. \tag{12}$$

Since it links two important, fundamental, atomic constants and can be measured with high precision, the Faraday has played an important and continuing role in the determination of the fundamental constants.

In his first adjustment of the fundamental constants, Birge [19] used the Faraday as an auxiliary constant. He deduced Avogadro's number from the ratio F/e using the value obtained for the charge of the electron from the oil-drop experiment. In later work, a reliable value for Avogadro's number became available from an absolute determination of the atomic spacing in, and the density of, calcite. The Faraday was then used by Birge [24] to obtain a value for the electronic charge differing by 0.6 percent from Millikan's oil-drop value. The error in oil-drop work was traced to an incorrect value used for the viscosity of air.

In modern determinations of the fundamental constants, adjustments are made by least squares procedures to give a best set of values for the fundamental constants that are compatible with the values of precisely measured atomic constants, such as the Rydberg constant and the fine structure constant. The Faraday has been used as an input datum in most of the adjustments that have been published to date. Together with the magnetic moment of the proton in nuclear magnetons, μ_p/μ_n, it strongly influences the adjusted value of Avogado's number.

Further work on the Faraday is highly desirable for the field of the fundamental constants. Confirmation of the new silver value, even at the 10 ppm level, would be sufficient to resolve the discrepancies between the various published values of μ_p/μ_n, which amount to 80 ppm.

8. The Future of the Faraday

The value formerly accepted for the Faraday was based on the electrolysis of silver salts in experiments carried out early in the century. By 1955 [25], it seemed clear from an examination of data on the fundamental constants that the iodine result of Vinal and Bates was, after all, the more reliable one. The latest determination, depending on the silver dissolution method, has eclipsed all the older work and the value of the Faraday remains based on the use of a silver coulometer, albeit of improved design. The comments of Washburn and Bates [13] therefore remain as appropriate today as when they were written in 1912 and they indicate clearly the direction that future work on the Faraday should take:

"*When we consider the amount of care and labour expended and the variety of reactions deemed necessary to employ in order to fix the atomic weight of a single element, it is certainly remarkable that, in the case of such an important and universal constant as the*

Faraday, scarcely any serious consideration has ever been given to an electrochemical reaction except that which occurs in the silver coulometer.'

The first priority is a redetermination of the iodine Faraday to confirm the agreement with the new silver value obtained by Zielen and Hamer from reworking of the original data of 1914. Such work is already in progress at the NBS and has recently been started at the NPL. Having due regard for uncertainties in the electrical units and impurities in the materials, any discrepancy that may remain between the two results must be sought, in the author's view, in the electrode reactions. Evidence that the silver and iodine electrode reactions are reversible has only been substantiated with uncertainties of about 20 ppm [11, 13]. Secondary electrode reactions are widespread in coulometry. It is conceivable that they should occur to the extent of a few parts per million and they have been invoked previously to explain differences between the various forms of coulometer. Only when agreement is achieved with coulometers employing different electrode reactions and conditions, can there be confidence that the underlying electrochemistry is understood and that the electrochemical Faraday is reliable.

9. References

[1] Lord Rayleigh and Mrs. H. Sidgwick, Phil. Trans. **175**, 411 (1884).
[2] Richards, T. W., and Heimrod, G. W., Proc. Am. Acad. Sci. **37**, 415 (1902).
[3] Kohlrausch, F., and Kohlrausch, W., Wied. Ann. **27**, 1 (1884).
[4] Guthe, K. E., Bull. Bur. Std. **1**, 21 (1904).
[5] Smith, F. E., Mather, T., and Lowry, T. M., Phil. Trans. **A207**, 545 (1908).
[6] Vigoureux, P., Metrologia **1**, 3 (1965).
[7] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375, (1969).
[8] Rosa, E. B., and Vinal, G. W., Bull. Bur. Std. **9**, 151 (1913).
[9] Rosa, E. B., Vinal, G. W., and McDaniel, A. S., Bull. Bur. Std. **9**, 209 (1913).
[10] Rosa, E. B., Vinal, G. W., and McDaniel, A. S., Bull. Bur. Std. **9**, 493 (1913).
[11] Rosa, E. B., Vinal, G. W., and McDaniel, A. S., Bull. Bur. Std. **10**, 475 (1914).
[12] Scott, A. F., Nat. Bur. Stand. (U.S.), Circ. 524, 1 (1953).
[13] Washburn, E. W., and Bates, S. J., J. Am. Chem. Soc. **34**, 1341 and 1515 (1912).
[14] Vinal, G. W., and Bates, S. J., Bull. Bur. Std. **10**, 425 (1914).
[15] Vigoureux, P., N.P.L. Coll. Res. **24**, 81 (1932).
[16] Craig, D. N., and Hoffman, J. I., Nat. Bur. Stand. (U.S.), Circ. 524, 13 (1953).
[17] Craig, D. N., Hoffman, J. I., Law, C. A., and Hamer, W. J., J. Res. Nat. Bur. Stand. (U.S.), **64A** (Phys. and Chem.) No. 5, 381–402, (Sept.–Oct. 1960).
[18] Ayrton, W. E., Mather, T., and Smith, F. E., Phil. Trans. **A207**, 463 (1908).
[19] Birge, R. T., Rev. Mod. Phys. **1**, 1 (1929).
[20] DuMond, J. W. M., and Cohen, E. R., Rev. Mod. Phys. **25**, 691 (1953).
[21] Sommer, H., and Hipple, J. A., Nat. Bur. Stand. (U.S.) Circ. 524, 21 (1953).
[22] Zielen, A. J., J. Electroanal. Chem. **18**, 89 (1968).
[23] Hamer, W. J., J. Res. Nat. Bur. Stand. (U.S.), **72A** (Phys. and Chem.), No. 4, 435–439 (July–Aug. 1968).
[24] Birge, R. T., Rev. Mod. Phys. **13**, 233 (1941).
[25] Cohen, E. R., DuMond, J. W. M., Layton, T. W., and Rollett, J. S., Rev. Mod. Phys. **27**, 363 (1955).

DISCUSSION

A. H. Cook: Can you not avoid the difficulties of chemical analysis by tagging your iodine radioactively?

A. Horsfield: This is an interesting suggestion. It would introduce complications in this experiment for, as soon as the coulometer current is started, ions throughout the bulk of the solution migrate towards the electrodes under the potential gradient between them. Untagged iodide ions in the vicinity of the anode will be discharged to iodine. Radioactive iodine, introduced around the cathode at the beginning of the experiment, will be reduced to iodide ions which will then start migrating towards the anode. You will therefore end up with a distribution of radioactive iodine throughout the solution. However, it might be possible to design a coulometer with the cathode compartment sufficiently well defined for the loss of radioactive iodine to be measured with the required precision.

J. Terrien: You asked a question about the name of units. There is being introduced a new base unit in the SI, the mole, and I think the mole is to be used in that case. For instance, the Faraday is the number of coulombs per mole—per mole of monovalent ions, of course.

A. Horsfield: What do you do about trivalent ions? For example, one mole of trivalent ions is then three equivalents.

J. Terrien: When you speak of the mole you have to specify the particle you deal with. So that the faraday is defined by means of monovalent ions; otherwise, the electric charge is three times larger for a trivalent.

A. Horsfield: All right. I think that is more awkward than using the equivalent.

U. Stille: I think there is a very easy solution. You see, the name of the quantity for which the mole should be the unit is I think in English quantity of substance. No, amount of substanee. Then in the special case that you have to do in electrochemistry with ions of different charge number—

A. Horsfield: Valence.

U. Stille: Oh, you say valency. Yes.

A. Horsfield: Yes.

U. Stille: Of different valency. Then you have to introduce only a new quantity, the "equivalent amount of substance," which is the product of the "n", amount of substance, and "z", the charge number of the ion, and there you are. You can use the mole all you like.

(*Applause*)

A. Horsfield: Very good. There will be a great many chemists who will be delighted if this suggestion is adopted. (*Laughter*)

Determination of the Faraday by the Iodine Coulometer

V. E. Bower

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

A determination of the Faraday by means of the iodine coulometer has been undertaken at the National Bureau of Standards. A progress report is presented. The article describes the iodine coulometer in use at present, the electrical equipment, and also details the chemical analysis of iodine by arsenious acid. The iodine generated in the coulometer can at present, by the method described, be measured with a precision of about 14 ppm (standard deviation of a single observation).

Key words: Coulometry; Faraday; iodine coulometer.

1. Introduction

In order to make an accurate determination of the Faraday by means of the iodine coulometer three conditions must be fulfilled. The first is that iodine must be generated or consumed in the coulometer with no side reactions. There is evidence that such conditions may be found [1, 2, 3, 4, 5]. The second

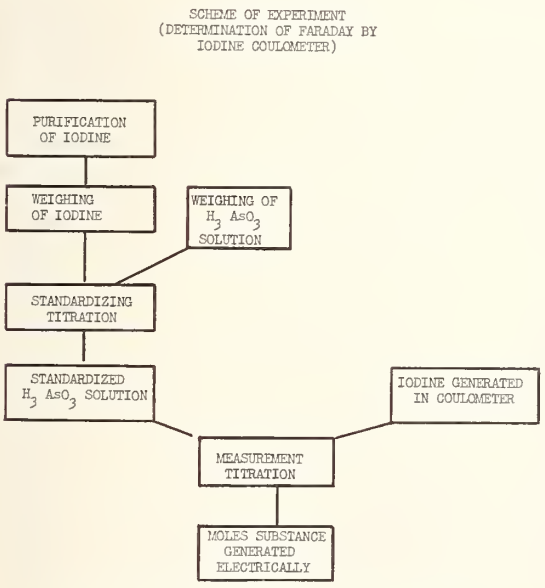


FIGURE 1. *Block diagram of the experiment.*
The diagram refers only to analysis of iodine generated at the anode.

is that we must be able to analyse the iodine generated or consumed in the coulometer with high precision. The third is that the required physical quantities, current and time, must be measured with high precision, modern notions demanding at least the level of parts per million. Needless to say, the third condition is more easily met than the first two.

The qualities of iodine which give it promise in a chemical determination of a fundamental constant are its ease of preparation in highly purified form in comparison to most other chemicals, the fact that it has one valence, and the fact that it is mononuclidic with an atomic mass known to one part per million or perhaps better.

This progress report will consist of a description of the techniques which we are using at the present time in our effort to measure the Faraday. In its essence, the method consists in standardizing a solution of arsenious acid with a highly purified sample of iodine and in using this same solution to measure the iodine generated or consumed in the coulometer. A block diagram of the experiment appears in figure 1. The description which we shall give here will take up the coulometer, the method of purifying the standard iodine, the method of keeping and weighing out the standard solution, and the method of titrating, with arsenious acid, the pure iodine sample and the iodine generated in the coulometer. This report covers work related to the anode reaction alone. The cathode reaction has not been studied to date.

2. The Coulometer

The present configuration of the coulometer (fig. 2) is a pair of cylinders with conical bottoms drained by tubes with stopcocks. The anode and cathode are of platinum sheet shaped to match the bottoms of the compartments and having holes at the apices of the cones for drainage. The electrical connections to the electrodes are platinum wires sealed into soft glass tubing. The drainage tubes are of sufficient length that they will reach down to the surface of the arsenious acid in the titrating flask below in order to reduce the danger of loss of material by splashing.

About 50 ml of a 10 percent solution of potassium iodide are run into both electrode compartments. By means of a separatory funnel a 50 percent solution of potassium iodide is introduced to a depth

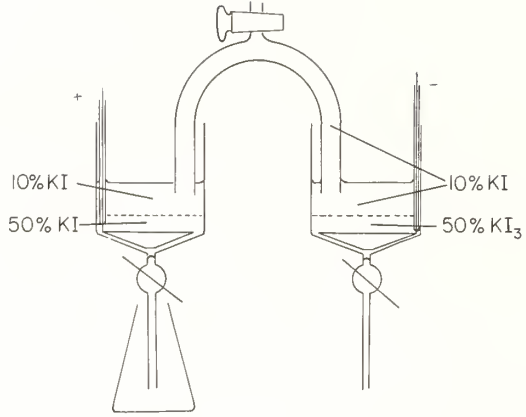


FIGURE 2. *The present design of the iodine coulometer used in the experiment.*

sufficient to cover the electrode under the 10 percent KI solution in the anode compartment. Similarly a 50 percent solution of iodine in the form of triiodide is placed in a layer under the 10 percent solution in the cathode compartment. After the layers are formed, some of the 10 percent solution is carefully drawn up to fill a bridging tube.

When current passes through the coulometer, iodide ion forms iodine (or what is the same thing, triiodide ion) in the anode compartment according to the reaction:

$$3I^- \rightarrow I_3^- + 2e \qquad (1)$$

and the reverse reaction occurs in the cathode compartment. The colored triiodide complex remains almost completely in the 50 percent iodide solution. The current is stopped before there is much diffusion of triiodide ion into the 10 percent solution.

At the end of the determination the bridge solution is drained, the stopcocks are opened and the solution in the anode compartment is allowed to run down into a flask containing a predetermined amount of arsenious acid solution. The clear 10 percent iodide solution rinses the triiodide ion into the tube. The titration is completed in a manner described later.

3. The Purification and Weighing of Standard Iodine

The apparatus which we are using for the purification of iodine is simply a long cruciform glass tube with a smaller entry tube for a stream of air which has been dried in a tall tower of silica gel (fig. 3). The tube arm without ground glass joint is placed in a small beaker which is held against the end of the open arm by a gentle spring. The arms with ground-glass joints are for access and manipulation. A tape heater which is made to cover about one third of the length of the tube is wound outside. Air at a controlled rate is admitted from the drying tower. A porcelain or glass boat charged with iodine is introduced into the upstream end and heating is applied.

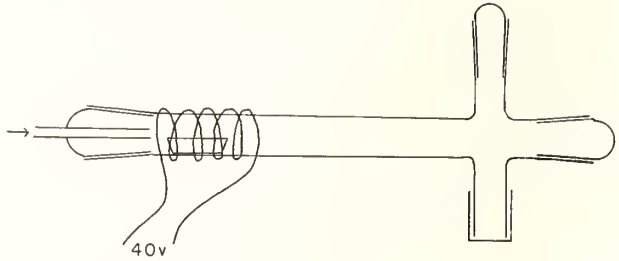


FIGURE 3. *Distillation tube for purification of iodine.*

The vaporized iodine is carried downstream beyond the heated zone where it forms slender and delicate crystals on the walls of the tube. When the boat is empty the heater is moved downstream thus driving the iodine toward the beaker where the sample finally arrives after three or four such zonal distillations. The residue which remains in the tube at the end of the distillation is knocked down into the beaker with a glass hoe pushed through the capped access tubes. Since each charge consists of iodine already distilled in the manner just described, every sample ultimately used has effectively been distilled about six times in air. Until it is used the iodine remains in the dry air stream. The beaker fits the distillation tube arm very closely. The iodine sample is thus protected from the contamination of the moisture in the laboratory air.

Wishing to weigh a volatile substance to a few parts per million and not wishing to be prodigal with material so laboriously prepared, we sought as our weighing container for iodine a small light, well-closed vessel which could be weighed and then dropped into the titration flask. The container is a simple glass cup 25 mm in outside diameter and 25 mm high. The lip of the cup is cut-off squarely with a diamond wheel and is ground flat on the side of the wheel. A 25 mm round microscope cover glass serves as top and is clamped to the cup by a gently sprung inconel clip. The clip is so bent that it is possible to assemble the cup without touching the edge of the cover glass. None of the vessels now in use loses more than one microgram per minute of iodine. A single-pan constant-load air-damped microbalance, with which one may make one weighing per minute, is used. One should be able to compute the sample weight at the very moment of immersion in the titrating solution to two micrograms or so. During weighings the cup is manipulated with gloved hands. Assembly and disassembly of the cup with lintless rayon gloves was tried many times and in no trial did the tare change more than 15 micrograms. The sample weight is usually 3–5 g.

4. Stabilizing and Weighing the Standard Solution

The solution used in the titration must be stored in such a way that the titer not change or that it change slowly enough to be useful. The solution used

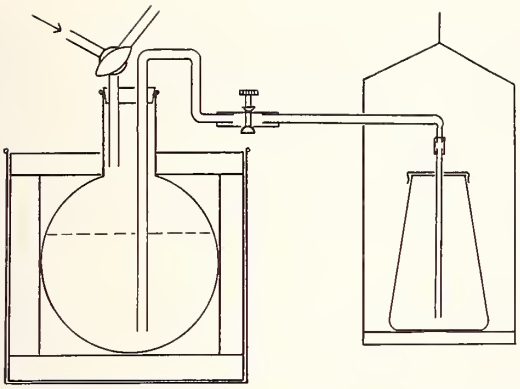


FIGURE 4. *Storage flask and delivery system for arsenious acid.*

in the present study is kept in a five-liter flask. The flask is enclosed in a sheet copper box lagged with blocks of polystyrene foam (fig. 4). The flask is fitted outside the lagging with a delivery tube for the solution and a three-way stopcock for the control of the argon under which the solution is kept. Argon introduced into the flask is bubbled through an identical stock solution to avoid changes in water vapor over the stock solution.

The room in which the analytical work is done is controlled to ±1.5 °C. The heat capacity of the whole system is large. The temperature of the solution should therefore vary only slowly.

The solution is delivered to a flask resting directly on a balance pan by a glass tube which terminates in a separable nozzle. Control of flow of the liquid is by a hosecock on a short piece of rubber tubing in the middle of the delivery tube. The strong alkalinity of the stock solution requires avoidance of greased stopcocks. Teflon stopcocks were found not to be sufficiently staunch over long intervals. The rubber tubing had been repeatedly boiled in NaOH solution and thoroughly leached in distilled water before being installed.

The flask into which the solution is delivered has always been wiped prior to weighing with a lint-free damp towel and allowed to come to equilibrium with the humidity of the room for about 30 min. The mouth of the flask is covered with a polyethylene film having a hole just large enough freely to pass the nozzle of the delivery tube. During the period of equilibrium, before the weighing, the humidity is adjusted inside the flask by purging with a jet of air which has been bubbled through a sample of stock solution.

In weighing the solution one clears the full volume of the delivery tube of stale solution. One then quickly attaches a new nozzle and places the equilibrated flask on the pan, tares it, runs in solution and takes the weight. The small size of the hole in the cover of the flask, the control of humidity inside the flask and the high concentration of the solution conspire to keep the rate of weight loss sufficiently low that one can make a good weighing.

5. The Titration

The quantity of iodine generated in the coulometer is determined by titration with a standardized solution of arsenious acid. The conditions for the accurate standardization of arsenious acid by iodine were set forth by Washburn [6] about 60 years ago. The improvements that the years have brought have been in the development of instruments for detecting the endpoint [7, 8] and for monitoring the hydrogen ion concentration [9]. The reaction is:

$$H_3AsO_3+I_3^-+H_2O\rightleftharpoons H_3AsO_4+2H^++3I^-. \quad (2)$$

From this the mass-action constant:

$$K=\frac{(H_3AsO_4)\,(H^+)^2(I^-)^3}{(H_3AsO_3)\,(I_3^-)}$$

$$=\frac{(H_3AsO_4)}{(H_3AsO_3)}\cdot\frac{(I^-)}{(I_3^-)}\,(I^-)^2(H^+)^2 \quad (3)$$

may be formed (parentheses indicate activity).

From not very exact values [10, 11] of this constant, Washburn computed the hydrogen-ion concentration required for a precise determination. Stripped of its details, his calculation ran as follows. If K is about 10^{-1} and we wish a precision of 1 ppm in the titration then the mol ratios $(H_3AsO_4)/(H_3AsO_3)$ and I^-/I_3^- will each be about 10^6. The conditions of the experiment are such that $(I^-)^2$ is about 10. We can solve for (H^+) and we obtain 10^{-7}. Today we should say that the pH be 7 at the endpoint. Furthermore we can monitor this condition by means of a glass electrode.

A more precise statement of the effect of variation of pH on the endpoint can be derived from some measurements of Marinenko and Taylor here at the Bureau. They made a precise coulometric assay of the arsenic trioxide standard sample issued by NBS. They found that a change in the pH of about $2\frac{1}{4}$ units brought about a change of 0.01 percent in the assay.

This means that if we seek a precision of 1 ppm in the titration we must control the pH to ±0.02 unit, a discrimination which lies well within the capability of modern pH meters and glass electrodes.

The titration for standardizing the arsenious acid is carried out as follows. The solid iodine is weighed in its cup. Enough solution is weighed out in the flask in an amount sufficient to react with all but 2–4 milligrams of the iodine. To the solution is added enough dibasic potassium phosphate solution to adjust the pH to 7 at the end-point. Sufficient potassium iodide is added to dissolve the iodine as triodide ion and to reproduce conditions in the coulometer. The iodine, dish and all, is placed carefully in the solution and allowed to dissolve and react with the arsenious acid. The small excess of iodine is titrated with a solution prepared from the stock solution by a two hundred fold dilution.

The endpoint of the determination is made amperometrically by a method similar to that of Ramsay,

Farrington and Swift, and Knowles and Lowden [7, 8]. Two platinium electrodes are suspended in the solution, a 0.15 potential between them. Current is measured on a microammeter. Dilute arsenious acid is run into the solution in 0.1 ml increments and the current is recorded after each addition. The current is linear in the iodine concentration near the end point and one may easily extrapolate the line to determine the volume equivalent to zero triiodide ion content. The sensitivity of the method as determined by repetition of the experiment on pairs of aliquots taken near the endpoint is 0.01 ml of the dilute solution. This represents 8 μg of iodine or 2–4 ppm for the size of sample generally used in these measurements.

When we standardize the arsenious acid solution with freshly prepared solid iodine we obtain a precise value for the titer of the solution which drifts with time at the rate of about 25 ppm per day.

We may assume a generally linear relationship to hold for the whole set of measurements shown. We may also assume (following a statistical model of Mandel [12]) that the difference between two successive observed values of the weight ratio are the result of a linear increase plus departures from linear increase that are independently and normally (i.e. Gaussian) distributed about a mean of zero with variance proportional to the length of the interval. Such an analysis places the estimate of the standard deviation of a point interpolated between observations at a maximum of 14 ppm.

The dispersion on the titer of the solution would appear to be only slightly larger than would be expected from the solution balance we are using. A more sensitive balance is being installed. No data are yet available from this arrangement.

6. Physical Apparatus

Current through the coulometer is controlled by a stabilizer consisting of two photogalvanometers in series. The IR drop across a standard resistor balances against a mercury cell. The error signal drives a galvanometer the patterned light beam of which falls on photodiodes in the arms of a bridge. The unbalance of this bridge drives another similar galvanometer which illuminates a second photo diode bridge. The output of the second bridge controls the bias on the control element of the power transistor which controls current flow.

The stabilizer will control current through good resistors to 2–3 parts in 10 million for short periods of time and to a few parts per million in the course of a day.

The value of the current is determined by matching the IR drop across a standard resistor to the potential of a standard Weston cell, or by having the IR drop measured directly against one of the cells in the NBS primary group.

The time interval is measured by a electronic counter which counts every ten pulses of the NBS 1 kHz standard frequency mains. A run of three hours should therefore provide a discrimination of 1 ppm in the time.

7. Conclusion

A progress report dealing with the component experiments in the determination of Faraday constant has been presented. This study is currently under way at NBS.

8. References

[1] Washburn, E. W., and Bates, S. J., J. Am. Chem. Soc. **34,** 1341 (1912).
[2] Bates, S. J., and Vinal, G. W., Ibid. **36,** 916 (1914).
[3] Lingane, J. J., and Anson, F. C., Anal. Chem. Acta **16,** 165 (1957).
[4] Lingane, J. J., and Kennedy, J. H., Ibid. **15,** 465 (1956).
[5] Marinenko, G., and Taylor, J. K., Anal. Chem. **39,** 1569 (1967).
[6] Washburn, E. W., J. Am. Chem. Soc. **30,** 31 (1908).
[7] Ramsey, W. J., Farrington, P. S., and Swift, E. H., Anal. Chem. **22,** 232 (1950).
[8] Knowles, G., and Lowden, G. F., Analyst **78,** 159 (1953).
[9] Bates, R. G., The Determination of pH (John Wiley & Sons, Inc., New York, 1964).
[10] Luther, R., Z., Electrochem. **13,** 289 (1906).
[11] Roebuck, J. R., J. Phys. Chem. **9,** 727 (1905).
[12] Mandel, J., J. Am. Stat. Assoc. **52,** 552 (1957).

DISCUSSION

B. P. Kibble: I just wonder what corrections you make for buoyancy when weighing the iodine in view of the fact you have got iodine vapor present in the trapped volume of air; and also if this volume of air is truly trapped it might not be at the same density as the surrounding air when you actually do the weighing.

V. E. Bower: Well, I took account of that. And, of course, it can be accurately calculated. I know the volume of iodine vapor over the sample pretty well because the sample size is always nearly the same. I think I may have as much as 6 micrograms in the vapor in the cup. So it doesn't have too much of an effect on the correction.

Precision Determination of Atomic Masses and the Einstein Mass-Energy Relation

A. H. Wapstra

Instituut voor Kernphysisch Onderzoek, Amsterdam, The Netherlands

The conversion constant between atomic mass and energy units is calculated in a least squares adjustment of atomic mass doublets and nuclear reaction energies. The agreement with the same factor as derived from Einstein's relation is satisfactory and comparable in precision with, but not independent of the value as derived from measurement of the Compton wavelength of the electron. Measurements are proposed that allow an independent check of Einstein's mass-energy relation.

Key words: Atomic masses; mass-energy equivalence.

1. Introduction

Essentially, two types of measurements are used in determining atomic masses of pure nuclides: mass spectroscopic measurements, and determinations of nuclear reaction energies. Somewhat loosely speaking, the first group gives the masses in mass units, the second one in energy units. Thus, in combining the two groups to get best values of atomic masses, the mass-energy conversion factor enters. Or, conversely, the combination of the two groups gives a test of Einstein's mass-energy relation. The present paper will discuss the validity of this test and compare it with the best other available tests. It will also give the best present values for the atomic masses of the most fundamental nuclides.

2. Mass Spectroscopic Measurements

Strictly speaking, the quantity measured in a mass spectrometer is not mass but mass ratio. Usually the ratio is that of the mass difference between two ions with (nearly) the same mass number (a "mass doublet") and that of one of the two ions. This ratio, $\Delta M/M$, is equal to the ratio of two resistors $\Delta R/R$ in a measuring circuit [1] connected with the mass spectrometer. (This explains the importance of the contribution to this Conference of L. Julie, who describes an improved circuit of this type.) From these mass ratios, the actual masses can be derived if one of them is known. Since the precision of the mass ratios is two orders of magnitude better than the precision that can be reached in absolute measurements, it was decided in 1963 to adopt an atomic mass scale by taking the mass of one atom of ^{12}C (with nucleus and electron cloud in their ground states) equal to exactly $12u$. Then, the ratios mentioned can be used to determine atomic masses of other nuclei on this mass scale.

Consulting the original papers [1], however, one will not find values for $\Delta M/M$ but only values for ΔM itself. These are obtained by multiplying the experimental ratios with the mass of the ion which is the better known member of the doublet. This does not introduce a serious systematic error that could invalidate a later best squares calculation. In any case of importance, the "better known" mass in a doublet is known with a precision of the order of 0.1 ppm, whereas the precision of the doublet rarely exceeds 10 ppm. (One should realize that, whereas the ratio $\Delta M/M$ is roughly of the order of one part in a thousand, the accuracy of mass spectroscopic measurements may be as high as one part in 10^8.)

The precision usually quoted for mass doublet measurements is obtained from the differences between a number of measurements of the same doublet from their mean value. Though this is an acceptable procedure, the error estimate obtained this way should be considered a lower limit, since it does not take into account the possibility of systematic errors. A next approximation to the real precision is obtained by making a least squares adjustment involving a number of doublets measured by one group of authors under comparable conditions and thus overdetermining a number of atomic masses. The least squares procedure (the χ^2 test) then yields a consistency factor R_e/R_i with the property that, if all original errors are multiplied with this factor, the input data are statistically consistent. This factor is found to lie between 1.5 and about 5. In the considerations below, all data on precision will be given taking into account the appropriate consistency factor [3]. The above procedure only makes allowance for part of the systematic error mentioned. However, after having made this correction, we find that the doublet measurements agree very well with nuclear reaction measurements, which suggests that the remaining part of the systematic error is of minor importance.

3. Reaction Energies

In precision nuclear reaction energy measurements, the quantity measured primarily is the influence of

Table 1. *Precision in some types of nuclear reaction energy measurements*

Reaction	Q-value	Reference
	keV	
$^{12}C(d, p)^{13}C$	2722.3±0.3	[a]
$^{212}Bi(\alpha)^{208}Tl$	6090.06±0.08	[b]
$^{3}H(p, n)^{3}He$	−764.03±0.15	[c]
$^{3}H(p, n)^{3}He$	−763.77±0.08	[d]
$^{3}H(\beta^-)^{3}He$	18.60±0.10	[e]
$^{3}H(\beta^-)^{3}He$	18.70±0.06	[f]
$^{3}H(\beta^-)^{3}He$	18.57±0.08	[g]
$^{1}H(n, \gamma)^{2}D$	2224.61±0.07	[h]
$^{1}H(n, \gamma)^{2}D$	2224.67±0.05	[i]

[a] O'Donnel F. H., and Browne, C. P., Phys. Rev. **158**, 957 (1967).
[b] Grennberg B., and Rytz, A., Compt. Rend. **269B**, 652 (1969).
[c] Rytz, A., Staub, H. H., and Winkler, H., Helv. Phys. Acta **34**, 960 (1962).
[d] Salgo, R. C., Staub, H. H., Winkler, H., and Zamboni, F., Nucl. Phys. **53**, 457 (1964).
[e] Porter, F. T., Phys. Rev. **115**, 450 (1959).
[f] Salgo, R. C., and Staub, H. H., Nucl. Phys. **A138**, 417 (1969).
[g] Davis, R., and St. Pierre, C., Nucl. Phys. **A138**, 545 (1969).
[h] Greenwood, R. C., Black, W. W., Phys. Letters **21** 702 (1966).
[i] Taylor, H. W., Neff, N., King, J. D., Phys. Letters **24**, 659 (1967).

an electric or magnetic field on the movement of charged particles, or the wavelength of electromagnetic radiations. A third type measurement, that of the velocity of particles (time-of-flight technique), has not yet been developed to comparable precision. All three types determine reaction energies in energy units. The precision reached in some of the best available measurements is shown in table 1. Even in the best reaction energy measurements of the type in which the energy of a secondary particle is measured, the precision is not better than about 100 ppm. That this is not due to the magnetic spectrometers used is shown by the measurement of the energy of the α-particles emitted by the radioactive nuclide ^{212}Bi (13 ppm). The obstacle to improvement of precision in the first kind of measurement is probably the influence of target thickness. In measurements where only the energy of the incoming particle is varied, observing thresholds or resonances in the yield of secondary particles, the available precision is slightly higher than in the other type of reaction measurements. Beta endpoints can seldom be measured to better than one part in 10^3, leading to high precision only where the transition energy is low, as in the case shown in table 1. Finally, gamma-ray energies have been measured with a precision of a few times 10 ppm in determining neutron binding energies with high precision. Thus, the best of the nuclear reaction measurements determine mass differences with nearly the same precision as mass doublet measurements. It is, therefore, appropriate on one hand, to use both kinds of data in determining best values for atomic masses of pure nuclides in a least squares adjustment and, on the other, to investigate to what extent the Einstein mass-energy relation is checked by these data.

4. Results

Carrying out the first part of this program requires knowledge of the mass-energy conversion factor:

$$1u = c^2/F \text{ eV} = 931\,481 \pm 5 \text{ keV} \qquad (1)$$

using for c and F the values calculated by Taylor et al. [4]. Our main calculation has been carried out using the value in (1) as a constant; some (preliminary) results are given in table 2, column IV. A somewhat better procedure would be treating c^2/F as an independent variable, with eq (1) as an additional input equation. For demonstration, this has been done for a limited set of data (all data on nuclides with mass number below 17); the results are compared in column III with results for the other kind of calculation on the same set of data presented in column I. The differences are only minor, as expected.

A logical method for checking the Einstein mass-energy relation appears to be that making an adjustment on the above data, treating c^2/F as a variable but *not* using eq (1) as an input equation. The result

$$c^2/F = (931\,512 \pm 48) \text{ keV}$$

if compared with eq (1), seems to check Einstein's relation with a precision of about 50 ppm (standard deviation). The set of atomic masses obtained in this way is shown for comparison in column II of table 2. The best other checks on Einstein's mass-energy relation are obtained from comparison of the energy of annihilation radiation with the mass of the electron. This comparison is simplified by the circumstance that Taylor et al. [4] chose not to use these data in their calculation of atomic constants from experimental data, mainly since the two best determinations of the annihilation energy do not agree very well: they are, respectively,

(ref. [5]: annihilation in H_2O) $\quad 1+(7\pm38)\times10^{-6}$

(ref. [6]: annihilation in Ta) $\quad 1-(56\pm16)\times10^{-6}$

larger than the electron mass recalculated in energy units using eq (1). The reportedly more accurate of these two values would point to a slight deviation but should probably not be taken quite that seriously. Together, they indicate that Einstein's relation is valid within 40 ppm, only slightly better than the value derived from atomic mass data as mentioned above.

It is necessary, however, to consider the last data somewhat more critically. In an earlier mass calculation [2], it was pointed out that the most important reaction energies were proportional to a calibration energy, that of the α-particles emitted by ^{210}Po. Thus, the precision of c^2/F determined from reaction and mass doublet data can never be higher than the precision in the ^{210}Po α-energy. This is no longer true. As discussed elsewhere [7], a value for c^2/F has already been determined with rather high precision

TABLE 2. *Mass excesses (M-A) for some nuclides as calculated:*
I not adjusting c^2/F;
II adjusting c^2/F too, no input equation;
III adjusting c^2/F, input equation (1); all three using only data on nuclides with $A<17$;
IV Provisional values of final adjustment.

	I	II	III	IV	
n	8 071.54±.07	8 071.73±.34	8 071.52±.07	8 071.54±.07	keV
	8 665.28±.06	8 665.19±.11	8 665.27±.07	8 665.27±.06	μu
H	7 289.05±.06	7 289.28±.37	7 289.04±.06	7 289.04±.05	keV
	7 825.23±.04	7 825.21±.05	7 825.22±.05	7 825.21±.04	μu
D	13 135.96±.11	13 136.40±.70	13 135.92±.12	13 135.90±.10	keV
	14 102.23±.09	14 102.10±.10	14 102.10±.10	14 102.21±.08	μu
T	14 950.02±.18	14 950.50±.80	14 949.90±.20	14 949.99±.17	keV
	16 049.73±.18	16 049.68±.19	16 049.68±.19	16 049.70±.16	μu
^{3}He	14 931.37±.18	14 931.90±.80	14 931.30±.20	14 931.33±.17	keV
	16 029.71±.17	16 029.70±.19	16 029.69±.19	16 029.67±.16	μu
^{4}He	2 424.71±.28	2 424.80±.40	2 424.80±.40	2 424.93±.26	keV
	2 603.00±.30	2 603.10±.40	2 603.10±.40	2 603.30±.27	μu
^{16}O	−4 736.38±.23	−4 736.53±.35	−4 736.37±.26	−4 736.61±.19	keV
	−5 084.78±.24	−5 084.77±.26	−5 084.77±.26	−5 085.10±.20	μu

from two simple combinations of the data as given in table 3. Attention is drawn here to the fact that probably without very much trouble the second of these two determinations can be improved by a factor of 3 or 4, by a better measurement of the $N(n, \gamma)$ capture gamma-ray energies, combined with an accurate determination of the ^{14}NH–^{15}N doublet.

The agreement between the two values in table 3 and eq (1) is satisfactory. Yet, deeper analysis of the data on the $H(n, \gamma)$ reaction show that their use in a check on Einstein's mass-energy relation is open

TABLE 3. *Separate determinations of* c^2/F

$^1H_2-{}^2D$	$=1548.20\pm0.07$ μu [a,b]
$^1H(n,\gamma)^2D$	$=2224.65\pm0.04$ keV [c]
$^3T(\beta^-)^3He$	$=18.64\pm0.04$ keV [c]
$^3T(p,n)^3He$	$=763.84\pm0.08$ keV [c]
1u = (931 514±80) keV	
$^{14}NH_2-O$	$=23809.25\pm0.3$ μu [d,e]
$^{15}NH-O$	$=13019.26\pm1.0$ μu [f]
$^{14}N(n,\gamma)^{15}N$	$=10833.1\pm0.5$ keV [g,h,i,j]
1u = (931 475±100) keV	

[a] Stevens, C. M., and Moreland, P. E., Proc. Intern. Conf., Winnipeg, 673 (1967).
[b] Johnson, W. H., Hudson, M. C., Britten, R. A., Kaiser, D. C., Proc. 3rd Intern. Conf. Atomic Masses, Winnipeg, (ed. R. C. Barber) p. 793 (1967).
[c] See table 1.
[d] Ref. [1].
[e] Smith, L. G., Phys. Rev. 111, 1606 (1958).
[f] Scolman, T. T., Quisenberry, K. S., Nier, A. O., Phys. Rev. 172, 1076 (1956).
[g] Carter, R. E., and Motz, H. T., Conf. Reactor Neutrons, Argonne, ANL-6797, p. 197 (1963).
[h] Bartholemew, G. A., Earle, E. D., and Gunye, M. R., Can. J. Phys. 44, 2111 (1966).
[i] Rasmussen, N. C., Hukai, Y., Inouye, T., and Orphan, V. J., MIT-NE-85, AFCRL-69-0071(1969).
[j] Greenwood, R. C., Phys. Letters 27B, 274 (1968).

to serious objections. The reaction-energy values given above have been derived from interpolation in a Ge(Li) spectrum of the escape peak of the capture radiation between the energies of the ^{60}Co gamma-ray lines of 1173.226±0.040 and 1332.482±0.046 keV [8]. Now, the energy of the capture radiation is obtained by adding twice the energy-equivalent of the electron mass to the energy of the escape peak mentioned. Furthermore, it appears that the energy values for the ^{60}Co lines mentioned are obtained, in final analysis, by comparison with the energy of annihilation radiation and therefore also from the energy equivalent of the electron mass. Since Einstein's relation is used in deriving these values, the comparison given above does not independently check this relation.

In the course of a rather extensive search for other (necessarily somewhat less accurate) measurements of the ^{60}Co or $H(n, \gamma)$ gamma rays that are free from the above objection, I found that the major part of the collection of high precision values for gamma-ray calibration energies is based, in final analysis, on the energy of the annihilation radiation. In particular, this is true for the $^{14}N(n, \gamma)$ capture radiations used in table 3. Thus, it has to be concluded that the precision of the check on the validity of Einstein's mass-energy relation from atomic mass data is not easily estimated, and certainly rather less than suggested by the comparison given above.

On the other hand, the somewhat more detailed analysis given here suggests that this situation could probably be improved considerably without too much trouble. This would require, on one hand, measurement of a few selected mass doublets like ^{12}CH–^{13}C and NH–^{15}N with the highest precision presently available. On the other, a somewhat complex program should be carried out of gamma-ray energy calibrations. In view of the fact that higher energy gamma ray lines can be measured with higher resolution than the annihilation radiation, one should expect a better check of Einstein's relation than obtained from the latter radiation. Also, such a

program would be valuable in itself since it would certainly help to improve the collection of gamma ray calibration energies.

5. References

[1] See, e.g., Quisenberry, R. S., Scolman, T. T., and Nier, A. O., Phys. Rev. **102**, 1071 (1958).
[2] Everling, F., König, L. A., Mattauch, J. H. E., and Wapstra, A. H., Nucl. Physics **25**, 177 (1961).
[3] Wapstra, A. H., and Gove, N. B., Nuclear Data Tables (1971).
[4] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[5] Knowles, J. W., Can. J. Phys. **40**, 257 (1962).
[6] Knowles, J. W., Proc. 2d Intern. Conf. Atomic Masses, Vienna (ed. W. H. Johnson), p. 113 (1964).
[7] Wapstra, A. H., Proc. 3d Intern. Conf. Atomic Masses, Winnipeg (ed. R. C. Barber), p. 831 (1967).
[8] Murray, G., Graham, R. S., and Geiger, J. S., Nucl. Phys. **63**, 353 (1965).
[9] Marion, J. B., Nuclear Data **4**, 301 (1968).

DISCUSSION

A. Rytz: I wanted to make a few short remarks regarding the alpha energy measurement. As you could guess from what Professor Wapstra said, we have set up at the BIPM an absolute spectrometer for measuring the energy of alpha particles from radioactive substances. Some results have been published—just the ones of ^{212}Bi. There are some more recent results available as to ^{242}Cm and ^{244}Cm, just preliminary ones. I hope that before the end of the year we shall have available about ten more measurements and published with comparable accuracy. As you could see, the imprecision we can get—it was a purely statistical precision which was on the slide—is of the order of 1 or 2 parts in 10^5. The total inaccuracy might be—we haven't yet worked out this well enough—but it might be of the order of 2 or not more than 3 times this value.

Then there is another remark I wanted to make just about this figure of 80 electron volts. I think this shouldn't be used to that extent in a mass calculation because there is an uncertainty due to what happens in the electron shells of the atom and which is not known to better than 100 or 200 electron volts.

A. H. Wapstra: Yes, I completely agree to that. Especially it would lead to serious compounding of systematic errors for a chain of such alpha decays, but as you know there are only very few alpha decays known to better than half a kilovolt. Thus, for most purposes it does not yet matter. I hope it will matter in about four years from now.

E. R. Cohen: In your specific examples you pointed out that you are really measuring gamma ray energies relative to the mass of the electron and so that you are not getting a measurement of the kilovolt/mass unit conversion factor. Does this really follow through in all the other energies that are used in the total mass system? Or is this unique to the few specific examples that you have identified in the end?

A. H. Wapstra: It is indeed true for these specific examples, but there are several more than mentioned in my talk. I would offhand estimate that if you take them out, the remaining precision in the conversion factor would be of the order of about 100 to 150 parts per million. I feel quite sure that the precision could be improved very considerably if a determined effort would be started. In that case, I think one ought to study a couple of special cases. One could, for instance, compare the Q-value of the reaction ^{13}C$(p, \gamma)^{14}$N with the mass doublet ^{13}CH–^{14}N, giving the same quantity directly in energy units and in mass units. I feel sure that both values can be measured now with a precision of about 20 parts per million, which would indeed give a very nice check on the conversion factor here.

E. R. Cohen: You have to be careful about how you calibrate your gamma ray energies, and make sure that you are not just converting them back to the annihilation radiation.

A. H. Wapstra: Yes, but this can be done.

E. R. Cohen: If you were serious in really trying to extract the value of the Faraday from all of this, you'd have to be very careful to identify how much of your energy scale for each reaction was really energy and how much was electron mass.

A. H. Wapstra: Yes. I completely agree. I said it was a big effort, and the main part will be to get an energy scale that is free from annihilation radiation.

W. C. Sauder: I'd just like to point out something that is going to happen tomorrow. Dr. Deslattes and I are collaborating on an experiment which will measure the annihilation radiation to a part per million, and in the process we are also going to make some measurements of nuclear gamma ray energies to the same level. So this ought to contribute to it.

An Improved Version of the Nier Circuit for $\Delta m/m$ Determinations

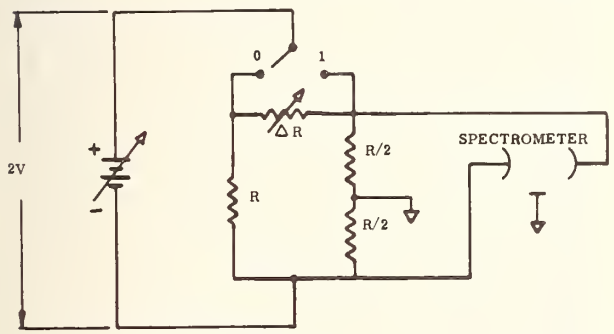
Loebe Julie

Julie Research Laboratories, Inc., New York, N. Y. 10023

In certain types of mass spectrographs $\Delta m/m$ is accurately related to $\Delta V/V$. Using the circuit configuration first described by Nier $\Delta V/V$ can be determined by the ratio $\Delta R/R$, where ΔR is a calibrateable rheostat and R represents the resistance of a set of fixed arms. This paper describes an improvement of the Nier circuit, using a recently developed potentiometer circuit in place of the conventional ΔR rheostat and fixed resistance arms. In the revised circuit $\Delta V/V$ is read directly on the dials of an accurate, high resolution Kelvin-Varley divider. Since it is believed that the best available Kelvin-Varley divider has an accuracy and resolution about one order of magnitude better than the best available rheostat the new circuit significantly extends the potential accuracy of $\Delta m/m$ determinations. This paper also describes a modification of the standard Nier battery circuit which eliminates the problems caused by battery potential shifts and drift during $\Delta m/m$ measurements.

Key words: Kelvin-Varley divider; mass spectrometry; Nier circuit.

The last one and one half decades have seen major improvements in the measurement of mass based on improvements in technology. These include new types of mass spectrometers, and new types of potentiometer circuits such as the basic $\Delta V/V$ circuit designed by Nier [1]. An additional major factor in making possible accurate inputs to the mass spectrometer is the development within the past 15 years of new types of ultrastable resistors suitable for use in instrumentation [2]. Most $\Delta V/V$ potentiometer circuits whether of the Nier or of other later designs [3], use resistors of the highest stability grades available in order to obtain extreme precision of electrical input to the deflection electrodes of the spectrometer.



$$V_0 = 2V\left[\frac{R/2}{R+\Delta R}\right] \quad , \quad V_1 = \frac{2V}{2} = V$$

$$\frac{\Delta V}{V} = \frac{V_1 - V_0}{V_0} = \frac{V - VR/[R+\Delta R]}{VR/[R+\Delta R]} = \frac{\Delta R}{R}$$

FIGURE 1. *Nier mass spectrometer circuit.*

Figure 1 shows the Nier circuit which is basic to present day mass spectrometry. In the circuit a voltage source is used which is typically in the range of 1000 to 3000 V, depending on the design of the spectrometer. In most cases this voltage source consists of a stack of suitable batteries, in series with a variable resistor (not shown) for voltage adjustment. The circuit generally uses four sets of carefully manufactured and adjusted resistors to establish a precise $\Delta V/V$ ratio. As shown in figure 1, the voltage applied to the mass spectrometer deflection plates at the right can be changed by an amount ΔV by introducing a resistance ΔR connected between points 0 and 1 of a selector switch or relay. If the voltages applied to the deflection plates of the spectrometer are V_0 and V_1 for the 2 positions of the switch and the ratio of the difference voltage $[V_1 - V_0]$ to V_1 is the proportional change in voltage applied to the spectrometer $\Delta V/V$ then $\Delta V/V = \Delta R/R$ as shown.

Since, in the Nier circuit, $\Delta V/V$ is equal to the proportional change $\Delta R/R$ of the resistors, the availability of high accuracy resistor sets has made possible extremely high precision in mass determination through precise control of the $\Delta R/R$ ratio. In the work reported by Stevens and Moreland at Argonne National Laboratories [4] a precise $\Delta R/R$ is achieved using a precision binary resistor set type BMR-105 for the R branch and strings of precision resistors type NB-1 for the R and $R/2$ branches. $\Delta R/R$ precision was of the order of $\pm 0.01\ \Omega$ at a ΔR value of 500 Ω and $\pm 0.1\ \Omega$ at a ΔR value of 70,000 Ω. This precision, which approaches 1 part per million for high values of ΔR was obtained by in situ intercomparison measurement and calibration of the ΔR and R resistors after they were mounted in a special constant temperature enclosure, to eliminate effects of temperature, at Argonne.

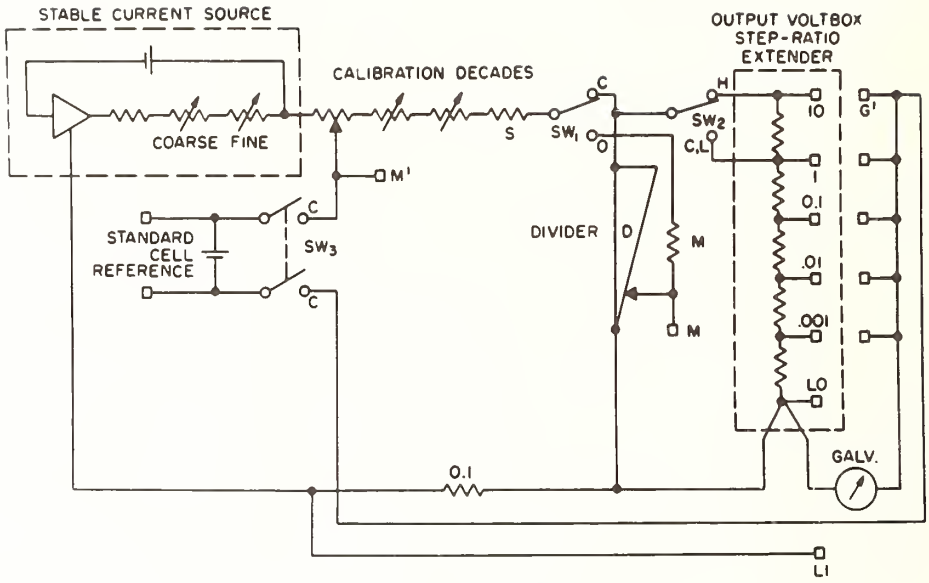


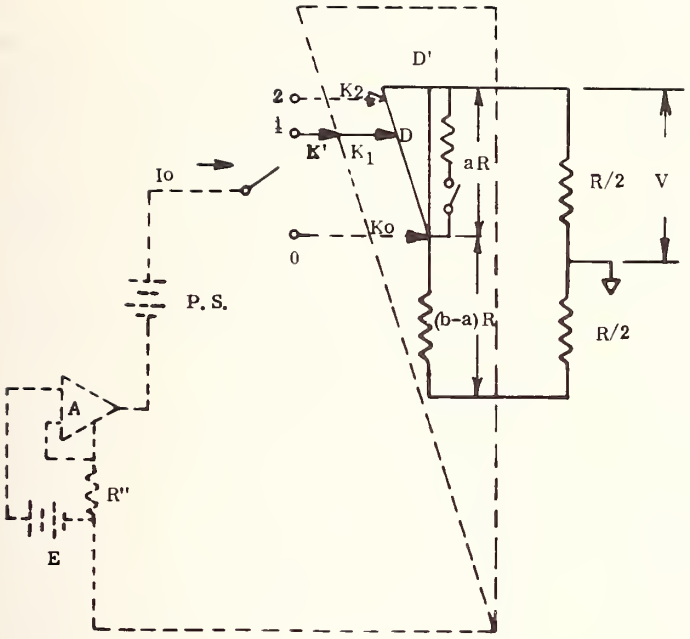
FIGURE 2. *Julie element potentiometer circuit.*

To conclude this brief description of the characteristics of the Nier circuit it should be pointed out that contact resistance variations in the switch or relay introduce errors into the $\Delta V/V$ system as do drifts or instabilities of the voltage supply during a run. In the referenced work by Stevens and Moreland a mercury relay was used to minimize the switch or relay resistance variations and a variable speed motor was used to provide a compensating voltage to adjust for drift of the battery supply.

The present paper describes a modification of the Nier mass spectrometer circuit using a Julie element potentiometer circuit shown in figure 2. This basic potentiometer uses a stable, regulated current source to deliver a fixed value of operating current to the terminals of a precise transfer ratio device D (typically a resistive divider) and connects the remaining terminals of the transfer ratio device or divider to an output circuit which may include a fixed load with suitable taps. As shown in papers recently published on potentiometers of this new type [5, 6], the output delivered to the load circuit is of the form I_0kZ where I_0 is the value of source current, k is the transfer ratio of the divider and Z is the effective impedance of the load circuit in parallel with the output impedance of the transfer ratio device. In theory, the accuracy of this type of potentiometer can be made extremely high over wide ranges of output voltage, by using current sources, transfer ratio dividers and load impedances of 1 ppm quality. Taylor, Parker, Langenberg, and Denenstein reporting on significant advances in Josephson Junction e/h determination have given experimental confirmation of the theoretically predicted accuracy of the Julie element potentiometer [7]. The reported uncertainty of 3 ppm in measurements at the 1 mV level shows the actual performance of the lowest range of this potentiometer.

The new $\Delta V/V$ circuit for mass spectrometers uses the basic potentiometer design of figure 2 with the modifications shown in figure 3. The circuit uses a transfer ratio device or divider D, with load resistors $R/2$ to drive the electrometer plates, and two scale factor resistances aR and $(b-a)R$ connected to the divider. To analyze the circuit simply, we may consider the combination of divider D and the scale factor resistors to be an equivalent divider D' which has a ratio k' and an output impedance R'. As shown in the potentiometer references previously cited the output voltage will be equal to $I_0k'Z/2$ where Z is the parallel combination of R' and R. k', R', and Z are easily determined as shown at the right of figure 3 giving the electrode potential V shown in eq (5). Letting V_0 be represented by the fixed term at the left of (5) and ΔV by the variable term containing k on the right we see that this circuit can deliver a potential, V_0, to the mass spectrometer electrodes, which is adjustable by varying I_0. It also delivers a separately adjustable fractional voltage $\Delta V/V_0$ which is precisely linear with divider ratio k, and which has a scaling factor of $c=a/(b-a)$.

Note that the scaling or range constant c for the new circuit is the ratio of the effective output resistance of the divider (which may be provided with switchable shunt resistors) to the resistance of the fixed set of resistors connected to the low side of the divider. The $\Delta V/V$ range constant does not depend on the output resistance value R. Thus in this circuit it is possible to adjust the full scale value of $\Delta V/V$ for different ranges of mass doublets, independently of the center tapped output resistor R which supplies potential to the spectrometer electrodes. It is also possible, for a given divider resistance, to choose a value for the resistor string at the bottom of the divider and values for the set of shunt resistors which will give decade ranges for $\Delta V/V$. Thus mass doublet



$$k' = \frac{(b-a)R + kaR}{bR} = \frac{(b-a)+ka}{b} \quad (2)$$

$$R' = bR \quad (3)$$

$$R'//R = \frac{b}{1+b} R \quad (4)$$

$$V = \frac{1}{2} I_0 k' \frac{b}{1+b} R$$

$$= \frac{1}{2} I_0 R \left[\frac{(b-a)+ka}{1+b}\right] = V_0 + \Delta V \quad (5)$$

$$\frac{\Delta V}{V_0} = \frac{a}{b-a} k = ck \quad (6)$$

FIGURE 3. *New mass spectrometer circuit.*

differences could be read on ranges of 0.01, 0.001, 0.0001 full scale with the full divider resolution on each range.

As shown in the figure the transfer ratio device may be used with a switch or relay operating between positions 0 and 1 in the usual way. The possibility of using a double divider with an extra switch position is also indicated. Use of the double divider would make it possible to adjust to two mass differences simultaneously, and to intercompare three mass values accurately. Since the author has no experience in spectrometry he is unable to say whether there is any practical value in this possibility.

Other practical features of the circuit are extremely useful. It is no longer necessary to be concerned about contact resistance variation of the switch since the circuit is driven from a virtually infinite impedance current regulating amplifier shown at the left of the diagram. A standard power supply or battery voltage is placed in series with the current source amplifier to establish the nominal high operating voltage for the system. This makes it possible to use stable low-voltage current sources, since the current source is not required to have high compliance voltage. In the new circuit the power supply voltage variations during a run do not affect operation unless they are so large as to carry the current source outside its compliance voltage range. It is worth noting that for the reasons given here it should be advantageous to use the author's proposed current source drive at the left of figure 3 to supply standard Nier as well as other basic types of $\Delta V/V$ potentiometers.

The accuracy of the new $\Delta V/V$ potentiometer depends on the ratio accuracy (linearity) of the divider and on the accuracy of range factor "c" which is determined by an intercomparison of aR with $(b-a)R$. Ratio calibration of the divider is carried out using a simple k/n ratio standard previously described [8]. Figure 4 shows experimental data of the extremely high accuracy obtainable with a new type of double divider [9, 10]. The data shows ratio accuracy with a worst case error of 3 parts in 10^8 and

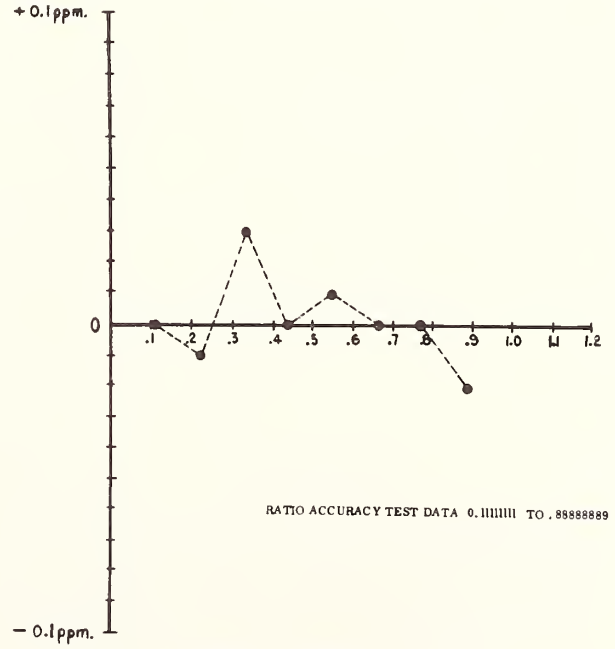


FIGURE 4. *Experimental data on accuracy.*

an rms error of 1.4 parts in 10^8. The aR versus $(b-a)R$ intercomparison is easily carried out by bootstrapping techniques involving either series-parallel laddering or calibration against the divider to accuracies of the order of a few parts in 10^7.

It is possible to use virtually any of the many types of dividers currently available with the Julie element $\Delta V/V$ circuit shown in figure 3. Depending on the spectrometry application involved these may range from low-cost multiturn slide-wire potentiometers, or simple 3 or 4 dial Kelvin-Varley dividers, to ultra-precise 7 or 8 place instruments. It also becomes possible to make use of data logging and programmable types of dividers and divider interfaces which are now available to accuracies of 1 ppm [11], so that mass spectrometer $\Delta V/V$ circuits may be operated under computer or program control and results data-logged automatically.

Conclusion

The proposed circuit appears to be extremely flexible and useful addition to the technology available for mass spectrometry where a $\Delta V/V$ potentiometer is required. It is based on a proven potentiometer design using ratio devices which are generally lower in cost and simpler to calibrate than rheostat devices of comparable accuracy. It has been described in a form suitable for fundamental research at the highest accuracy level although it is easily modified for use at lower accuracies with simpler and less expensive dividers and current sources. It also affords the opportunity for automatic printout, programming or computer control of the $\Delta V/V$ input to the electrodes of the mass spectrometer.

References

[1] Nier, A. O., Nuclear Masses and Their Determination (Pergamon Press, Inc., London, 1957), p. 185.

[2] Julie Research Laboratories Circular PRECISION, Vol. 1, No. 1, Resistors of Very High Absolute Accuracy and Stability, December 15, 1958.

[3] Bishop, R. L., and Barber, R. C., A high precision ratio potentiometer, Rev. Sci. Instr. **41**, 327 (1970).

[4] Stevens, C. M., and Moreland, P. E., Mass Measurements with the Argonne 100 in. Mass Spectrometer, Proceedings of the Third International Conference on Atomic Masses, R. C. Barber, Ed. (University of Manitoba Press, Winnipeg, 1968), p. 673.

[5] U.S. Patent 3,454,877; Julie, Loebe, A Universal Potentiometer for the Range From One Nanovolt to Ten Volts; IEEE Trans. Instr. Meas. **IM-16**, 187 (1967).

[6] Abramowitz, Abraham, Principles, Current Practice, and Applications of Potentiometers Based on Poggendorff's Second Method, Rev. Sci. Instr. **38**, 898 (1967).

[7] Taylor, B. N., Parker, W. H., Langenberg, D. N., and Denenstein, A., On the Use of the AC Josephson Effect to Maintain Standards of Electromotive Force, Metrologia **4**, 89 (1967).

[8] Julie Research Laboratories Circular PRECISION, Vol. 1, No. 2, A Precision Voltage Divider With One Part-Per-Million Absolute Accuracy, December 29, 1958. Vol. **3**, No. 5, Establishing the High Absolute Accuracy of the VDR-106 Primary Standard Voltage Divider, May 27, 1960. Vol. **4**, No. 1, Establishing Ratios to One Part in Ten Million, June, 1961.

[9] Dean, L. W., and Ludwig, D. R., The Julie-Dean Double Potentiometer, Measurements and Data **4**, No. 1, 74 (January–February, 1970).

[10] Julie, Loebe, A New 8-place DC Potentiometer for High Volume Testing of Linear Components, Electronic Instrument Digest **6**, No. 10, 46 (October, 1970).

[11] Julie, Loebe, A Fully Automated DC System Which Can Measure or Generate, V, I, R, and K to 1 Ppm Accuracy, Conference on Precision Electromagnetic Measurement, Boulder, Colorado, June 1–4, 1970.

PROTON MAGNETIC MOMENT

Methods of Measuring the Magnetic Moment of the Proton in Terms of the Nuclear Magneton

B. W. Petley

Quantum Metrology Division, National Physical Laboratory, Teddington, Middlesex, England

The paper reviews the methods used prior to 1970 to measure the magnetic moment of the proton in nuclear magnetons.

Key words: Atomic constant; magnetic moment; μ_p/μ_n; precise measurement; proton.

1. Introduction

The measurements of the magnetic moment of the proton μ_p in terms of the nuclear magneton μ_n, are closely related to the prewar measurements of the charge to mass ratio of the proton. The latter experiments required a measurement of the proton cyclotron frequency f_c in a uniform magnetic flux B, the two being related to the proton charge to mass ratio e/M_p by

$$f_c = B(e/2\pi M_p). \tag{1}$$

Protons have spin as well as charge and so have a magnetic moment. The energy change $2\mu_p B$ associated with the transition from the magnetic moment being aligned parallel to being antiparallel to the magnetic flux requires the absorption of a photon of energy hf_s, where h is the Planck constant and f_s is

the proton spin precession frequency. Thus

$$2\mu_p B = hf_s. \tag{2}$$

If the gyromagnetic ratio of the proton, $\gamma_p = 4\pi\mu_p/h$ is known the value of e/M_p may be obtained by using eq (2) to deduce the magnetic flux. Unfortunately however γ_p is not yet known with sufficient precision to be regarded as an auxiliary constant so that the values of e/M_p measured in this way would be subject to change.

Alvarez and Bloch (1940), [1] pointed out that if B is eliminated from eqs (1) and (2) by simple division then

$$f_s/f_c = \mu_p/(he/4\pi M_p). \tag{3}$$

The quantity in the denominator is the nuclear magneton μ_n, and this is the value expected for the proton magnetic moment from simple quantum theory involving the proton charge and angular momentum. The actual value for μ_p differs from μ_n by an amount that at present is not ealculable. Modern practice therefore is to properly regard the experiment as measuring μ_p/μ_n rather than e/M_p.

2. The Proton Spin Precession Frequency, f_s

A whole branch of physics has become based on techniques of nuclear magnetic resonance (nmr) and the measurement of f_s is therefore relatively straightforward [2]. The signal from the precessing protons is small and so a large proton density is required to observe the resonance. Although signals can be detected from protons in gaseous hydrogen which is at the order of atmospheric pressure, it is usual to observe the resonance signal from a sample of protons in water. Heavy water has also been used, and also liquid paraffin, with appropriate corrections. There is a diamagnetic correction for the shielding of the protons in the water and since this correction is about 26 ppm and the exact value has not yet been established [3], it is customary to publish the μ_p/μ_n results with and without this correction. There are other corrections for the sample shape etc., some of which have been mentioned in the conference review of the measurements of γ_p by Vigoureux.

In order to observe the spin precession frequency the water sample is placed inside a coil that is tuned to its resonant frequency and the magnetic field modulated locally by a small amount. The 'Q' of the tuned circuit changes each time that the conditions for nuclear magnetic resonance are satisfied. The nmr signal is normally observed by one of two methods in μ_p/μ_n determinations. The simplest method is the amplitude bridge method pioneered at the NBS by Thomas and Huntoon (1949) [4]. In their method the change in the circuit 'Q' may be detected by supplying a signal to the coil from a radio frequency generator and rectifying it with a diode, followed by low frequency amplification and phase-sensitive detection synchronous with the modulating frequency.

The change in 'Q' may also be detected by making the coil part of the tuned circuit of a marginal oscillator. This type of oscillator is operated in the condition where it is only just oscillating so that a change in the circuit 'Q', with the field modulated locally as before, results in a larger change in the level of the oscillator. This method was pioneered by Pound and Knight (1950) [5]. The second method is the more sensitive and gives a better signal to noise ratio by about a factor of ten. The amplitude bridge method has the advantage of simplicity. The signal from protons in water is sufficiently strong for the disadvantage of the poorer signal to noise ratio to be unimportant, especially if narrow-band detection techniques are employed.

The difficult part of the work lies in the measurement of the proton cyclotron frequency and the remainder of this review is devoted to the methods that have been evolved over the last twenty years for measuring this frequency.

3. The Proton Cyclotron Frequency

The measurement of f_c is limited in precision by two problems which are a direct result of the electrostatic field or electrostatic field gradients that are present in any vacuum apparatus. An electrostatic field E_s, causes the centre of curvature of the orbit to drift with a velocity $E_s \wedge B/|B|^2$, which is in a direction perpendicular to B. Thus in a device having potential variations of the order of a tenth of a volt in distances of the order of a centimeter and in a flux density of 1 tesla, the drift velocity is about 10 m/s. This drift velocity is generally mass independent and the ions drift to the walls, limiting the time during which the cyelotron frequency can be measured to the order of a millisecond. These fields are of importance where the cyclotron frequency is deduced by defining the orbit precisely since they represent a frequency shift.

An electrostatic field gradient $\partial E/\partial z$ in the magnetic flux direction (z) and a space charge density ρ, perturb the cyclotron frequency from that given by eq (1) to one typically of the form [6]

$$f_c^1 = f_c - [\rho/E_o - \partial E/\partial z]/B. \tag{4}$$

Failure to allow for the effects of this shift affected the pre-war e/M determinations and has been the major difficulty in the post-war measurements. (Note: although $\partial E/\partial z$ has been used the shift is fundamentally a result of the radial electric fields.)

The mass ratios of ions are known with sufficient precision [7] for them to be regarded as auxiliary constants. It is not necessary therefore to measure the cyclotron frequency of protons. This allows the electrostatic frequency shift to be eliminated by bringing different mass ions successively to resonance, provided that changing the resonant ion does not vary either the space charge or the electrostatic field gradients. Alternatively measurements may be made using the same ion at different magnetic flux densities and a similar assumption made about the

invariance of the bracketed term in eq (4). Both methods have been employed, the former being the most used.

A number of methods of observing the cyclotron frequency have been evolved over the years and these are conveniently divided according to whether the initial ion energy is low or high, or whether the frequency is observed by accelerating or decelerating the ions, or at constant energy (by a time of flight method).

4. Low Energy Methods: the Omegatron

The low energy methods of observing f_c are derived from that developed at the NBS by Sommer, Thomas, and Hipple (1951) [8, 9]. Since the device allowed the ion cyclotron angular frequency ω_c to be measured it was termed an 'omegatron'. As well as being of importance in the derivation of the fundamental atomic constants the omegatron became of considerable importance in vacuum physics and more than 200 papers have been published concerning the omegatron. The omegatron has contributed for example to the reliability of the repeater amplifiers in the trans-Atlantic telephone cable and to our knowledge of the gases in the upper atmosphere. The number of published papers may be taken as an indication that the simplicity of the omegatron is deceptive.

4.1. The Sommer, Thomas, and Hipple Determination

The operation of the omegatron may be understood from reference to figure 1. A beam of electrons, ~1 μA, 70 eV, moving in the magnetic flux direction, form thermal ions in the analysis region. The ions are then accelerated, by applying a radio-frequency field, until they hit the collector. The collected ion current is in the region of 10^{-12} A.

Ions arrive at the collector whatever their starting phase relative to the radio-frequency field. The frequency width of the resonance depends on the time of flight of the ions, and is independent of the ion mass. Sommer, Thomas, and Hipple increased the containment time for the ions and hence the resolution, by applying an electrostatic field which trapped

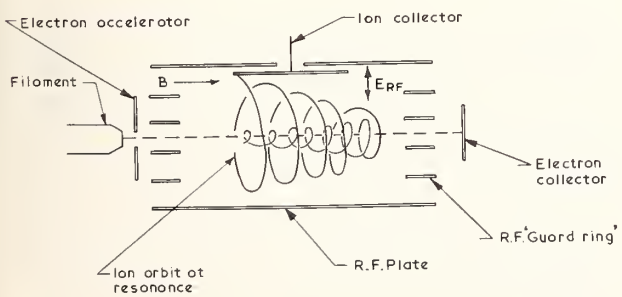


FIGURE 1. *The essential features of an omegatron.*

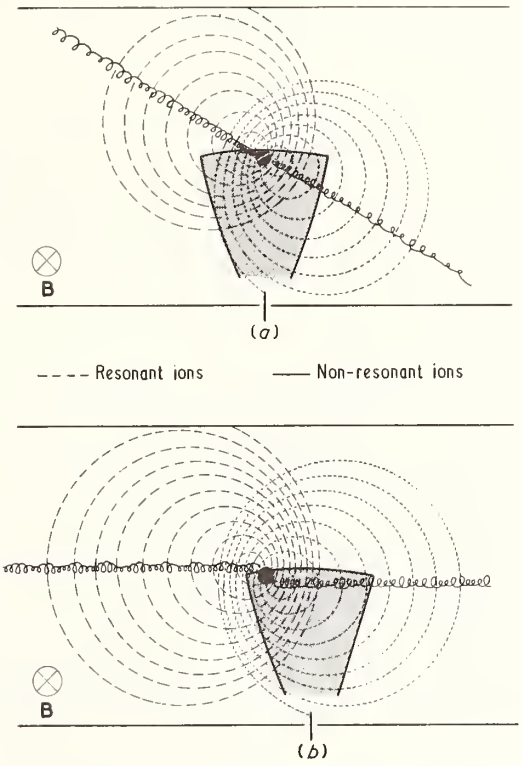


FIGURE 2. *The drift of ions in an omegatron (a) the rectangular type and (b) the quadrupole type (after Petley and Morris), J. Sci. Instr.* **42**, 492 (1965).

the ions in the magnetic field direction. The trapping field was required to overcome the thermal velocity of the ions which is in the region of 100 to 1000 m/s. Unfortunately the trapping field has two effects, (i) it renders the orbit unstable in the direction

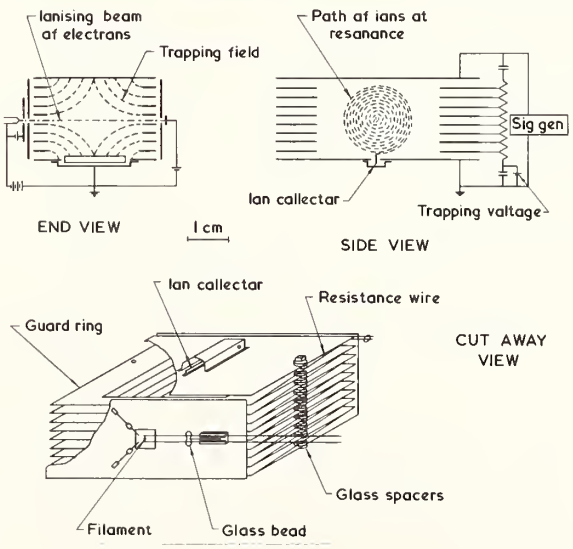


FIGURE 3. *The omegatron of Sommer, Thomas and Hipple (after Sommer, Thomas and Hipple, Phys. Rev.* **82**, 697 (1951).

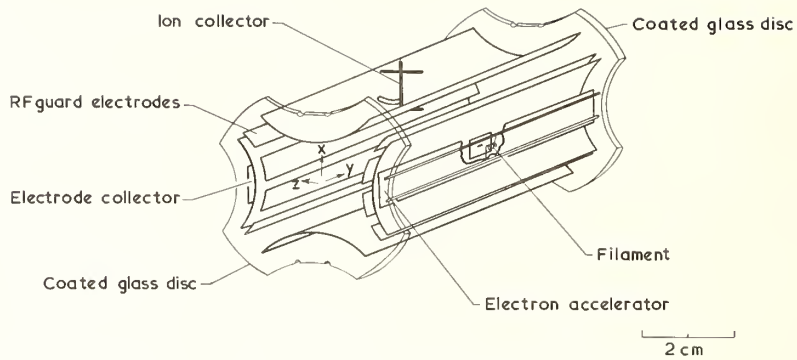


FIGURE 4. *A cut-away view of the quadrupole omegatron (after Petley and Morris NPL Report QU7 1969).*

perpendicular to the magnetic field, again limiting the ion lifetime and (ii) the trapping field gradient shifts the ion cyclotron frequency as in eq (4). The ions follow the path indicated in figure 2a. The centre of curvature of the orbit must remain within the region bounded by the parabolae if the ions are to hit the collector and not be lost by colliding with the surrounding electrodes. At a given magnetic flux density there is a limit to the narrowness of the resonance that is otherwise governed by the dimensions of the omegatron and the ion thermal energy [10].

Although the Sommer, Thomas, and Hipple Omegatron, figure 3, had guard rings to render the radio frequency field uniform it is clear from their paper [8], Sommer's Ph.D. Thesis [9] and their beautifully kept laboratory notebooks, [11] that the omegatron was operated with the guard electrodes grounded for rf, that is with a non-uniform rf field. Since the ions are effectively orbiting at a frequency that is midway between f_c and the applied radio frequency, any non-uniformities in the rf field, perturb the cyclotron frequency, as deduced from the centre of the observed resonance curve, by effectively adding to the expression in eq (4). Such non-uniform fields do not necessarily affect the final result if it can be established that all ions follow essentially the same path so that the electric fields are averaged in the same way.

In order to eliminate the electrostatic field shifts Sommer, Thomas, and Hipple used pairs of ions, H^+ and H_2^+, H^+ and D_2^+ and H^+ and H_2O^+. It will be observed that H^+ ions, protons, are common to all pairs and it is interesting that the procedure advocated by Breitenberg and Cohen [12] if valid in this case, (reference [8], figure 6) would lead to a result that was several ppm higher than the published value.

One of the difficulties with using protons arises from the fact that the proportion of protons relative to H_2^+ is pressure dependent [13, 14] because they are produced by a second order process (ionisation and dissociation). The possible disparity in the ion intensities might affect the invariance of eq (4) because the elimination of the bracketed term depends on the space charge term being independent of the mass of ion that is at resonance.

A further difficulty with using protons is that they are produced by breaking up a molecule and the results of Dunn and Kieffer [15], if valid in an omegatron, indicate that many of the protons are produced with more than an electron volt of initial energy. Since the initial ion energy determines the orbit drift velocity, the space charge term could become mass dependent. (The equality of the initial ion energies in an omegatron can be gauged quite simply; the protons would require a higher trapping voltage than the H_2^+ ions for maximum collected ion intensity if their initial energy were significantly higher). As a result of the difficulty of getting sufficient proton intensity, neither of the other two published omegatron determinations have used protons.

4.2. The Omegatron of Petley and Morris

The Sommer, Thomas, and Hipple omegatron was rectangular in shape and used plane electrodes. The trapping potential distribution cannot be a simple function of the space coordinates because of the boundary conditions and the equations of motion can only be solved in an approximate manner. The omegatron of Petley and Morris [16, 17, 18], figure 4, was an attempt to overcome this difficulty. The electrodes were made hyperbolic in shape and the glass end plates coated with a conducting film to simulate an infinitely long omegatron. The trapping potential distribution ϕ_T is simply

$$\phi_T = V_T\{k_o + k_1(z^2 - x^2)\}$$

where x and z are the rf and magnetic field directions respectively, k_o and k_1 depend on the omegatron dimensions and V_T is the applied trapping voltage.

The same electrodes are used in an omegatron to simulate a hyperbolic trapping potential distribution and a uniform radio-frequency field. The latter requirement was achieved in the quadrupole omegatron by splitting two of the hyperbolic plates into four sections. An electrolytic tank was used at the design stage to evaluate the best relative potentials for the rf guard plates in order to achieve an essentially uniform rf field over the active volume.

The equations of motion in a quadrupole omega-

tron may be solved exactly. The ions drift with a uniform velocity that depends only on their initial distance from the electric axis as shown in figure 2b. Owing to the construction tolerances the electric axis can differ slightly from the axis of the electron beam, and this results in a greater proportion of the ions drifting in one direction than in the other, (construction tolerances of better than the 0.1 mm achieved would otherwise be required). Two ion collectors were used, and the one giving the greater collected ion current used for the measurements.

Both the rectangular omegatron (STH) and the quadrupole omegatron (PM) were used at a flux density of about $0.5T$ so that they were operated in the orbit drift limited mode to give a narrow resonance. The resonance shape therefore departed from the rectangular or trapezoidal shape that is expected from simple theory. In this mode particularly the position of the centre of the resonance, which indicates f_c, varies with the height up the resonance. This variation is a direct result of the variation of the electrostatic and magnetic fields over the ion orbit.

Petley and Morris made measurements at three heights up the resonance using H_2^+, HD^+ and D_2^+ ions. Their results were independent of height to within 2 ppm. Thus although omegatron resonances are in practice slightly skew it appears that a consistent result may be obtained for μ_p/μ_n. Reference to the STH notebooks indicates that they made measurements at the resonance half-height.

There are three possible ways of reducing the data to one final value but it can be shown that provided the method of least squares is employed the same answer should be obtained within the rounding errors. In the PM measurement all three methods were used as a check, the rounding error being 0.3 ppm [18]. The major error in the PM measurement arose from the non-uniformity of the magnetic field which contributed 12 ppm.

4.3. The Boyne and Franken Omegatron

In both of the previous measurements the magnetic field from the filament current shifted the cyclotron frequency by a few ppm and was eliminated by taking readings with the filament current in both directions. In the Franken and Boyne determination (FB) [6] an alternating current was used to heat the electron filament and so the perturbing field was averaged out. The main magnetic field was also in the region of one tesla. In their measurement the electrostatic frequency shifts were eliminated by varying the magnetic flux density over the range 0.8 to $1.25T$.

The omegatron was of simple construction as shown in figure 5. Whereas in the previous two measurements the ions gained several hundred electron volts of energy while gaining enough energy to hit the collector, in the Boyne and Franken determination the resonance condition was observed by using marginal oscillator techniques. The advantages of this were a better signal to noise ratio and a much smaller change in the ion energy, only a fraction of

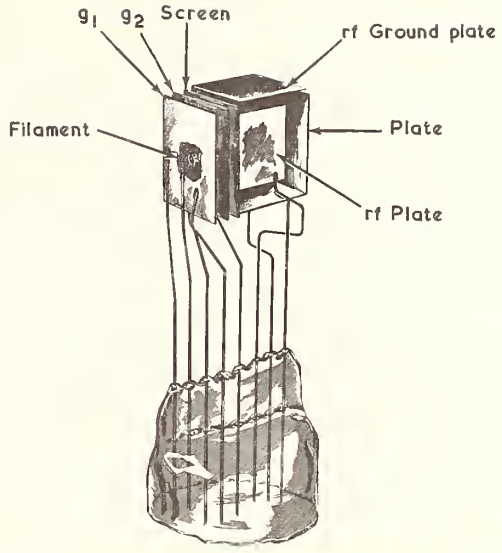


FIGURE 5. *The omegatron of Boyne and Franken (after Boyne and Franken, Phys. Rev.* **123**, *242 (1961).*

an electron volt. The short ion lifetime resulted in a broader resonance than in the other omegatrons but this could be split very precisely by using narrow-band detection techniques. The resonance was adjusted to symmetry by altering the potentials applied to the various electrodes. The radio frequency field was not uniform, but this was not important since the ions gained very little energy from it.

They used H_2^+ ions and used a D_2O sample to measure the nmr frequency. The latter frequency was converted to f_s for protons by using the known gyromagnetic ratio of D_2O relative to H_2O. The reasons for this were that the spin precession and cyclotron frequency differed by only about 16.5 percent which simplified the problem of measuring the frequency ratio. This was an advantage in the measurements since the shifts in the cyclotron frequency were eliminated by varying the magnetic field and so both f_s and f_c varied. Thus they used

$$f_c^1 = f_c(1 - \Gamma/B^2) \qquad (5)$$

which is an alternative form of eq (4) where Γ is essentially the electrostatic field term.

The inherent difficulty with this method lies in the assumption that the electrostatic field term does not change with magnetic field at least over the experimental range of magnetic fields. Boyne and Franken considered errors of this type very carefully and assigned a standard deviation of 18 ppm to their possible effect.

The methods where different mass ions are brought to resonance rely on the invariance of the space charge, independently of which mass of ion is at resonance and are therefore equally vulnerable. The electrostatic field and space charge effects of eq (4) exist in *all* methods and their elimination is a major problem in any measurement of the proton cyclotron frequency.

5. High Energy Methods: Acceleration

5.1. The Method of Bloch and Jeffries [19]

Bloch and Jeffries reported a measurement of μ_p/μ_n in 1950 which used a cyclotron in the decelerating mode. Their method had the advantage over the accelerating cyclotron that the magnetic field could be made uniform since the decelerating radio-frequency field provided focussing of the ions. Their apparatus is illustrated in figure 6. The diameter of the dees was 8.5 cm and the flux density about 0.5T.

The protons were generated outside the magnetic field and their path kept straight up to the instant of injection by using electrostatic fields to counteract the Lorentz force. These fields penetrate into the cyclotron through the injection slit and so perturb the cyclotron frequency. The number of revolutions that can be made by the ions in this method is limited because they must lose sufficient energy during the

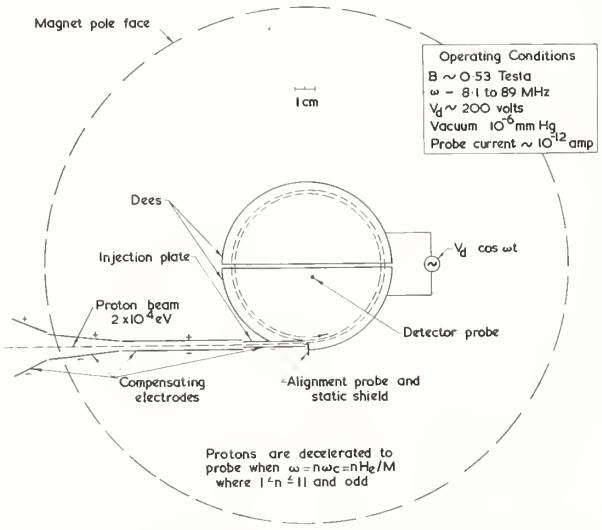


FIGURE 6. *The decelerating cyclotron of Bloch and Jeffries (after Bloch and Jeffries, Phys. Rev.* **80** *305 (1950).*

first revolution to clear the injection slit. This requirement corresponds to a limiting radio frequency voltage V_c which enters into the measurement of f_c. The value of V_c is deduced from measurements of the collected ion current as a function of frequency using higher decelerating voltages because at $V_d = V_c$ no ions can reach the ion collector. The number of revolutions, and hence the resolution, is increased slightly because transit time effects reduce the deceleration of the ions in crossing the dees at smaller radii.

The ions do not spend all of their time in the radio frequency field and so they may be decelerated by applying an harmonic of the cyclotron frequency. The use of an harmonic frequency results in a narrower resonance curve because only ions of a narrower band of starting phases reach the collector. The resolution is improved in proportion to the harmonic used. In practice the resolution cannot be increased

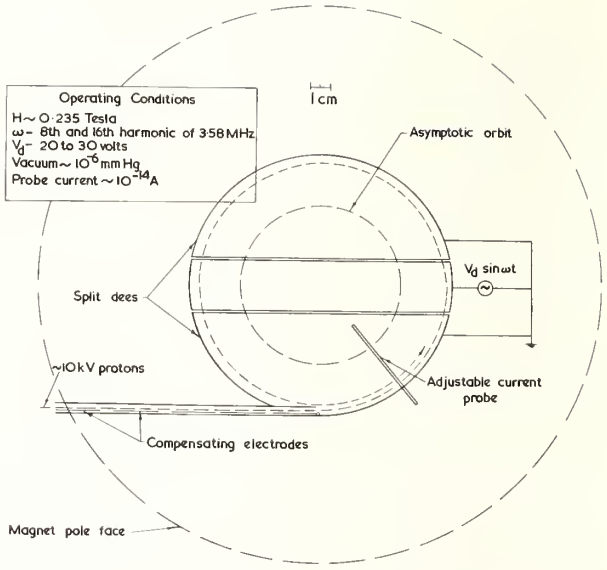


FIGURE 7. *The modified decelerating cyclotron of Sanders (after Sanders and Turberfield, Proc. Roy. (Lond.)* **A272**, *79, 1963).*

indefinitely owing to defocusing and other effects at the dee gap. Bloch and Jeffries used up to the ninth harmonic of f_c which enabled fractional linewidths of a few parts in 10^3 to be observed, even though the ions made only a few hundred revolutions. The voltages applied to the dees were in the region of 220 volts (peak) which is more than 200 times greater than the rf voltages applied in the omegatron determinations.

The use of a highly non-uniform radio frequency field means that the resonance curve is no longer symmetric about f_c. Ions only reach the collector when the applied radio frequency is below the cyclotron frequency, or below a corresponding harmonic of f_c. Thus it is the upper limit of the observed resonance curve that is required, and this is the frequency at which the collected current has fallen to zero. Unfortunately a detailed theory for the shape of the resonance is required in order to deduce this upper frequency limit correctly. The value of $(\mu_p/\mu_n)''$ deduced from the upper frequency limit varies with the rf voltage as

$$(\mu_p/\mu_n)'' = (\mu_p/\mu_n)\{1-(\Delta f_r/nf_r)[(V_d/V_c)^2-1]^{1/2}\} \tag{7}$$

where Δf_r is the resonance width and nf_r the applied radio frequency. A major contribution to the final error of 72 ppm in the Bloch and Jeffries determination is the uncertainty in estimating the upper frequency limit, the other errors amounting to about 15 ppm. Trigger [20] later repeated the determination with improved apparatus and made a more detailed analysis of the effects of the electric fields. His final standard deviation was 36 ppm and the precision was still largely limited by the uncertainties involved in deducing the cyclotron frequency from the observed resonance curve.

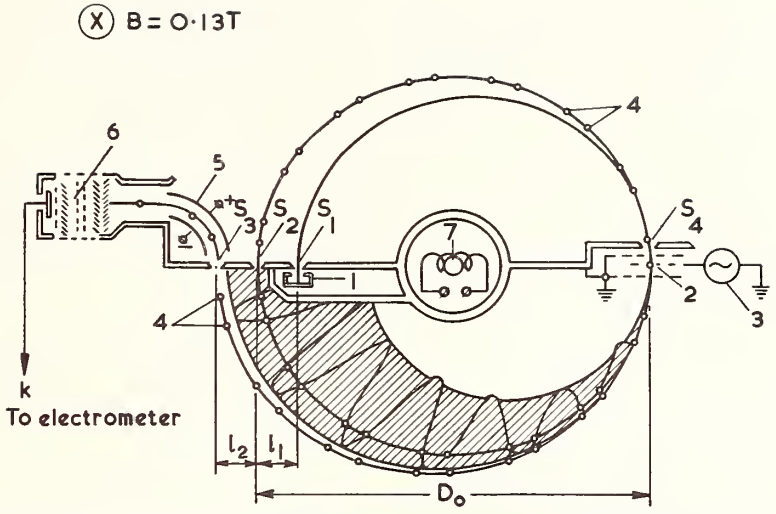


FIGURE 8. *The resonance mass spectrometer of Mamyrin and Frantsuzov (after Mamyzin and Frantsuzov, Sov. Phys. JETP* **21**, *274 (1965).*

5.2. The Decelerating Split-Dee Cyclotron of Sanders

The limited resolution resulting from the relatively small number of revolutions made by the protons in the above work was considerably improved in an ingenious modification by Sanders [21, 22, 23] figure 7. The dees of the cyclotron were separated by a centre electrode 3 cm wide, the dee diameter being 14 cm. The protons were therefore subjected to a decelerating rf field four times per revolution instead of twice. The resonances were narrower because there was a smaller range of starting phases for which protons reached the ion collector. The protons tended towards an asymptotic orbit whose diameter was such that the ions received no net deceleration per revolution.

The protons reached the asymptotic orbit after about 150 revolutions. The rf voltage was only about 35 V (peak). The thickness of the injector plate determines the decrease in energy during the first revolution and the values of V_d corresponding to two thicknesses of injector plate were 34.4 ± 0.7 and 27.4 ± 0.7 V.

Although the magnetic flux density was less than half that used by Bloch, Jeffries and Trigger, $0.235T$ compared with $0.53T$, the base width of the resonance was narrower and gave nearly a fourfold increase in the resolving power over that achieved by Trigger. The effects of the electric fields were again analysed very carefully in order to check correctness of the deduction of f_c from the observed resonance curves. Once again this error was the dominant one in the final error estimate, the latter corresponding to a standard deviation of 26 ppm.

6. High Energy Methods: Acceleration

The third high energy method is the Russian determination by Mamyrin and Frantsuzov [24, 25, 26]. In their method the ions made only two complete revolutions, traversing the radio-frequency field twice. However the resolution achieved by them was a further factor of about four better than that achieved by Sanders et al. This improvement in precision was largely a result of their use of a very much higher harmonic of f_c, in the region of the 200th harmonic. The apparatus of Mamyrin and Frantsuzov is illustrated in figure 8.

In their method the ions are generated in the magnet gap and have an initial energy of from 400 to 200 eV, depending on their mass. The energy is sufficient to give the appropriate mass ion a radius of curvature of 9.8 cm in a flux density, provided by a permanent magnet, of $0.13T$. The first half revolution serves to separate out ions of different mass. The ions of chosen mass pass through the slit S_1 where they are subjected to an accelerating rf voltage of up to 1 kV amplitude. The resulting spread in the ion energies depends on their arrival phase with respect to the rf field. The range of the ion trajectories lies within the shaded region in figure 8. Two short bunches of ions of a particular narrow band of energies (or starting phases) emerge from the slit S_2. If these ions arrive at the slit S_1 with the same phase as previously, they are accelerated again and pass through the slit S_3. This occurs if $F_{osc}=nf_c$. After passing through the slit S_3 the ions are electrostatically deflected onto the electron multiplier.

The use of the electron multiplier, which was shielded from the magnetic field and had a gain of about 10^4, enabled ion currents of 10^{-15} A to be measured with a good signal to noise ratio. The omegatron and inverse cyclotron methods have the disadvantage that they must have the ion collector connected directly to the electrometer and so are unable to take advantage of the high ion energy to improve the signal to noise ratio.

Since the ratio of l_1 to D_0 is 0.125 the ion energy is changed considerably. The velocity increment of the ions at the slit S_2 does not vary sinusoidally with time,

Table 1. *Summary of the pre-1970 measurements of* μ_p'/μ_n

Method	Flux density (T)	Ions	μ_p'/μ_n	Error (in ppm)
Omegatron				
Sommer, Thomas, Hipple (1951)	0.47	H^+, H_2^+, D_2^+, H_2O^+	2.792690(60)	22**
Boyne and Franken, (1961)	0.8–1.3	H_2^+	2.792701(73)	20
Petley and Morris, (1967)	0.47	H_2^+, HD^+, D_2^+	2.792746(52)	19
Decelerating Cyclotron				
Bloch, Jeffries, Trigger, (1956)	0.53, 0.7, 0.96	H^+	2.79267(10)	36
Sanders, Dellis, Turberfield, (1963)	0.24	H^+, H_2^+	2.792701(73)	26
Magnetic resonance mass spectrometer				
Mamyrin, Frantsuzov, (1965)	0.13	He^+, Ne^{++}, Ne^+	2.792794(17)	6.2
Nuclear Reaction Energies				
Marion, Winkler, (1967)	0.75	H^+ from $\begin{cases}^{27}Al(p,\gamma)^{28}Si \\ ^{7}Li(p,n)^{7}Be\end{cases}$	2.79260(13)	45

** The interpretation of this error is subject to debate, see text.

it also varies for the two traversals of the modulators and is different for the two pulses of ions.

The second velocity increment is therefore smaller than the first and so the distance l_2 is less than l_1 by about 10 percent. The position of the slit S_3 is adjusted so that at a particular oscillator frequency f_{res} both bunches of ions are received simultaneously.

For the above condition the relation between f_{res} and f_c is

$$f_{res}=f_c(n+k) \tag{8}$$

where k is much less than unity. The value of k was obtained by measuring f_{res} as a function of n and also calculated by solving a set of six transcendental equations which contained sine and cosine terms. Both methods were consistent with $k=-0.0028$. The term k arose from the fact that the acceleration of the ions while traversing the modulator modified their cyclotron period. The theoretical calculations of Mamyrin and Frantsuzov assumed that the ions traversed the modulator grid section during $\frac{3}{4}$ of an rf half-period. For $n=196$ the outer modulator grids were separated by 1.35 mm and for $n=96$ they were separated by 2.70 mm (the grid separation must be varied for operation at different harmonics of f_c).

Mamyrin and Frantsuzov were able to observe the frequency shifts due to the electrostatic field terms and the shifts were between -15 and 90 ppm on the Ne^+ ions and correspondingly less on the lower charge to mass ratios of the He^+ and Ne^{++} ions that were also brought to resonance. Changes are unfortunately required in some of the electrode potentials in order to bring the different mass ions to resonance and the extrapolation procedure only eliminates completely the electrostatic-field dependent frequency shifts that remain constant. The effects of all of the electrostatic field shifts and other effects that could change the cyclotron frequency were carefully analysed and it was concluded that the total systematic error did not exceed 5 ppm. This error was combined with the random error of 4 ppm to give an overall experimental error of 6.2 ppm.

7. High Energy Method: Time Flight

This determination of μ_p/μ_n [27, 28] arose out of discrepancies between workers at the University of Maryland and the University of Zurich in their determination of certain proton-nuclear reaction energies, namely the $^{27}Al(p,\gamma)^{28}Si$ resonance (992 keV) and $^{7}Li(p,n)^{7}Be$ neutron threshold (1881 keV).

In the Maryland experiment the velocity of the proton was measured by a time of flight method. They modulated the intensity of the proton beam by applying a large electric field perpendicular to the line of flight which caused the proton beam to move across an exit slit so modulating the beam intensity. The modulated beam then entered a resonant radio-frequency cavity. The entry of the proton pulse into the cavity resulted in the propagation of a pulsed electromagnetic wave down the cavity, thereby exciting it at its resonant frequency (about 72 MHz). A similar pulse was generated as the bunch of protons left the cavity. When the two waves were exactly out of phase there was a sharp null in the signal detected by a pick-up loop that was located near the exit slit, (the cavity 'Q' was 4000). The condition for a null signal was that

$$L/V_p=(2N-1)/2f_{res}, \qquad N=1, 2, 3 \tag{9}$$

where L (1.25 m) was the cavity length and f_{res} ($\simeq$72 MHz) was the resonant frequency. Their major experimental difficulty was that of measuring the cavity length L. This length is the electrical length and not the geometric length. The measured geometric length must be corrected to allow for the perturbation of the electric field by the apertures. This resulted in an uncertainty of 18 ppm.

The Zurich apparatus measured a quantity that is closely related to the beam velocity, namely the magnetic rigidity

$$B_p\rho'=M_pv_p/e\{1-(v_p/c)^2\}^{1/2} \tag{10}$$

where ρ' is the proton beam orbit radius in the magnetic flux B and M_p is the proton rest mass. If the magnetic field is measured by nmr techniques, yielding the spin precession frequency f_s, the equation reduces to

$$\mu_p/\mu_N=2\pi f_s\rho'\{1-(v_p/c)^2\}^{1/2}/v_p. \tag{11}$$

Therefore by using the Maryland data to deduce v_p the value of μ_p/μ_N may be obtained from two

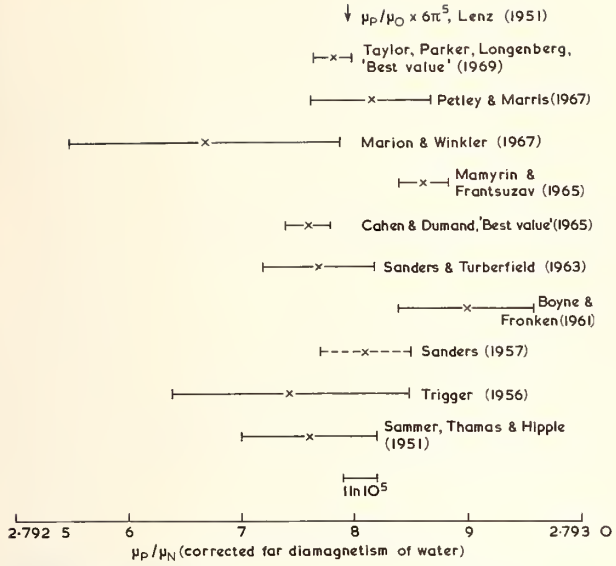


FIGURE 9. *Summary of the pre-1970 values of μ_p/μ_n without the correction for the diamagnetism of water.*

experiments. The precision obtained in the initial measurement was 44 ppm. The experiments suffered from the obvious disadvantage that the same protons did not pass through the two sets of apparatus and this disadvantage has been removed by shipping the Maryland velocity gauge to Zurich. The results of the more recent experiments are to be reported at this conference.

8. Results

The results of all of the determinations are summarised in table 1. With the exception of the Sommer, Thomas, and Hipple determination the values given are those in the Taylor, Parker, and Langenberg (1969) [3] review. The STH determination is controversial mainly through the interpretation that is to be placed on their published error. In the nineteen-fifties it was usual to publish results with the probable error. Generally the reviewers have tended to regard the error as a 2σ error. On the other hand the experimenters, myself included, have always regarded their error as being one standard deviation. The error quoted therefore is that given in the original paper.

The results are shown in figure 9. It is apparent that if the STH error is regarded as one standard deviation the results are statistically compatible with some intermediate mean value. The recommended best values of Cohen and DuMond (1965) [29] and of Taylor, Parker and Langenberg (1969) [3] have also been shown.

The determination of μ_p/μ_n limits for example, our knowledge of M_p/m_e, since

$$\mu_p/\mu_n = (\mu_p/\mu_B) \cdot (\mu_B/\mu_n) = (\mu_p/\mu_B)\,(M_p/m_e) \qquad (12)$$

where μ_p/μ_B is known to parts in 10^7. For interest I have shown the value of μ_p/μ_n corresponding to the suggestion by Lenz [30] that M_p/m_e is close to $6\pi^5$. It is interesting that this value remains compatible with the experimental evidence.

(Note. A similar combination has occurred earlier with the suggestion by Lewis and Adams (1914) [31] that

$$1/\alpha = 8\pi(8\pi^5/15)^{1/3} = 137.348.)$$

9. Acknowledgment

This work forms part of the research programme of the National Physical Laboratory.

10. References

[1] Alvarez, L. W., and Bloch, F., Phys. Rev. **57**, 111 (1940).
[2] Andrew, E. R., Nuclear Magnetic Resonance (Cambridge University Press), (1958).
[3] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[4] Thomas, H. A., and Huntoon, R. D., Rev. Sci. Instrum. **20**, 516 (1949).
[5] Pound, R. V., and Knight, W. D., Rev. Sci. Instrum. **21**, 219 (1950).
[6] Boyne, H. S., and Franken, P. A., Phys. Rev. **123**, 242 (1961).
[7] Mattauch, J. H. E., Thiele, W., and Wapstra, A. H., Nuclear Phys. **67**, 1 (1965).
[8] Sommer, H., Thomas, H. A., and Hipple, J. A., Phys. Rev. **82**, 697 (1951).
[9] Sommer, H., Ph. D. Thesis, Agricultural and Mechanical College of Texas (unpublished) (1950).
[10] Petley, B. W., and Morris, K., J. Phys. E. **1**, 417 (1968).
[11] See the note by Feystrom, D. O., Petley, B. W., and Taylor, B. N. (submitted to this conference) (1970).
[12] See discussion following reference (18).
[13] Buckingham, J. D., Thorne, P., and Wingate, H., J. Sci. Instrum. **42**, 737 (1965).
[14] Bajeau, G., and Comsa, G., Trans. 8th Symp. Vacuum Technology (Pergamon Oxford), 617 (1962).
[15] Dunn, G. H., and Kieffer, L. J., Phys. Rev. **132**, 2109 (1963).
[16] Petley, B. W., and Morris, K., J. Sci. Instrum. **42**, 492 (1965).
[17] Petley, B. W., and Morris, K., Proceedings of the Third International Conference on Atomic Masses, R. C. Barber, Ed. (University of Manitoba Press, Winnipeg), 461 (1968).
[18] Petley, B. W., and Morris, K., Nature **213**, 586 (1967).
[19] Bloch, F., and Jeffries, C. D., Phys. Rev. **80**, 305 (1950).
[20] Trigger, K. R., Bull. Amer. Phys. Soc. II **1**, 220 (1956).
[21] Sanders, J. H., Nuovo. Cim. Suppl. **6**, 242 (1957).
[22] Collington, D. J., Dellis, A. N., Sanders, J. H., and Turberfield, K. C., Phys. Rev. **99**, 1622 (1955).
[23] Sanders, J. H., and Turberfield, K. C., Proc. Roy. Soc. A **272**, 79 (1963).
[24] Mamyrin, B. A., and Frantsuzov, A. A., Dokl. Akad. Nauk, U.S.S.R. **159**, 777 (1964) [Soc. Phys.-Doklady **9**, 1082 (1965)].
[25] Mamyrin, B. A., and Frantsuzov, A. A., Zh. Eksp. Teor. F. 2, **48**, 416 (1965) [Sov. Phys.-JETP **21**, 274 (1965)].
[26] Mamyrin, B. A., and Frantsuzov, A. A., Proceedings of the Third International Conference on Atomic Masses, R. C. Barber, Ed. (University of Manitoba Press, Winnipeg), 427 (1968).
[27] Marion, J. B., Proceedings of the Third International Conference on Atomic Masses, R. C. Barber, Ed. (University of Manitoba Press, Winnipeg), 410 (1968).
[28] Marion, J. B., and Winkler, H., Phys. Rev. **156**, 1062 (1967).
[29] Cohen, E. R., and DuMond, J. W. M., Rev. Mod. Phys. **37**, 537 (1965).
[30] Lenz, F., Phys. Rev. **82**, 556 (1951).
[31] Lewis, G. N., and Adams, E. Q., Phys. Rev. (2) **3**, 92 (1914).

Magnetic Moment of the Proton in Nuclear Magnetons*

D. O. Fystrom

Wisconsin State University, La Crosse, Wis. 54601

A redetermination of μ_p'/μ_n using an omegatron similar to that of Sommer, Thomas, and Hipple is reported. Particular emphasis was placed on reducing rf field inhomogeneity and estimating the influence of surface potential variations.

Key words: Fundamental Constants; μ_p/μ_n; omegatron.

1. Introduction

One of the compelling reasons for continued interest in measurement of μ_p/μ_n is the bimodal distribution in its values. I report here on a redetermination of μ_p/μ_n using an omegatron as shown in figure 1. The omegatron used is similar to that of Sommer, Thomas, and Hipple (STH) [1]. The magnetic field was locked to the frequency of a crystal via a frequency offset phase lock loop controlled NMR oscillator as shown in figure 2. The NMR oscillator design was essentially that of Robinson [2]. A schematic of the oscillator stabilization technique used for measurement of the magnetic field is shown in figure 3. Both NMR oscillators were frequency swept.

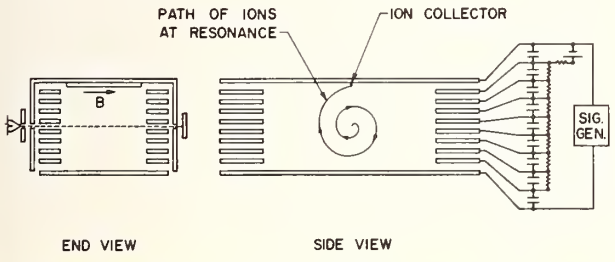


FIGURE 1. *Schematic of the rectangular charge collectiug omegatron used.*

2. Potential Divider and Lineshape

One of the points of concern at the time of the publication of STH was the asymmetric line shape. During preliminary explorations of the properties of the new omegatron it was noted that the leads coupling the guard rings to the end plates were much longer than those coupling between guard rings. I guessed that this might have significant influence on the divider impedance and therefore on the rf electric field uniformity thus causing the asymmetry of the observed lines. Tests showed that lineshapes could be modified drastically and the line center shifted by using impedance steps of unequal size.

2.1. Frequency Shift and Nonuniform RF Electric Field

A theory for the shift can be given in terms of one of the circularly polarized components of the rf electric field. Others [3] have shown that to a very good approximation the actual orbit in a plane polarized rf electric field is given by considering only one of the circular polarized components. This accelerating component may be represented by $(E_0/2)\exp[j(\omega t+\phi)]$ where ω is the rf angular frequency and ϕ is the phase. E_0 and ϕ may depend on position in the omegatron. The actual frequency over one orbit is $\omega+(d\phi/T)$ where $d\phi$ is the change of the average ϕ from one orbit to the next and T is the orbit period. A similar argument shows that an rf electric field which changes in direction with change in position may also lead to a frequency shift.

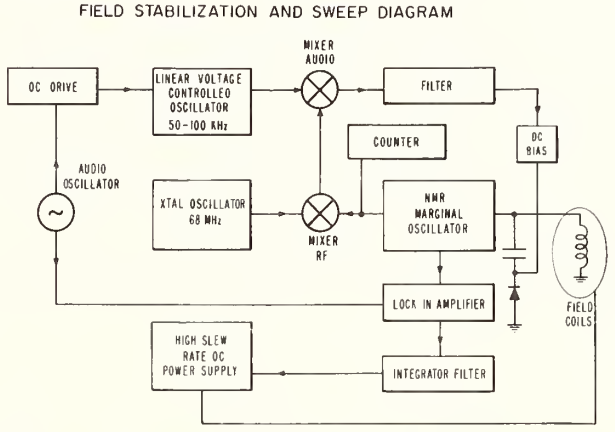


FIGURE 2. *Schematic of the circuitry for field stabilization and field sweep utilizing a frequency offset phase lock loop.* The frequency sweep was about 200 Hz.

* This work supported by NBS Laboratory, Astrophysics Division. JILA Building, University of Colorado Campus, Boulder, Colo. 80302.

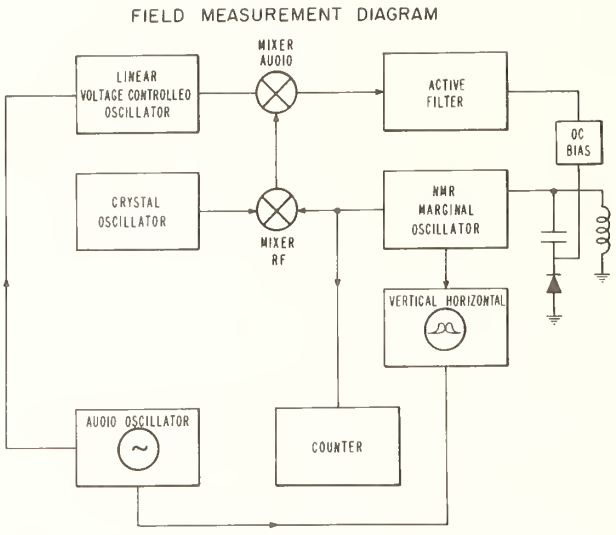


FIGURE 3. *Schematic of the circuitry for field measurement.*
It also utilizes a frequency offset phase lock loop.

2.2. Nearly Rectangular Lineshapes Observed

When the voltage divider was modified so that all steps were equal, line shapes such as those in figure 4 resulted for the omegatron under the condition of severe oxide coating. With increasing E_{rf} the lines become more rectangular until bumps in the wings of the line appeared at a threshold which was dependent on the trapping field value. These horns disappeared for low trapping voltage when the omegatron was gold plated. Proton lines of the type shown in figure 5 have asymmetries of about 0.5 ppm. I believe these lines can be thought of as the result of having the orbit motion dominated by the rf electric field and the magnetic field, with other fields taking a more minor role in determining the line shape. Shifts such as those related to orbit drift are reduced by going to larger rf electric fields. Special care must be taken to avoid increasing those shifts due to change of average phase with radius in the higher electric field.

3. Trapping Field Shift Correction

The relationship between the observed cyclotron frequency ω_0 and the electrostatic field free result ω_c for an ion of mass m is $\omega_0=\omega_c\{1-[E(r)m/B^2er]\}$ where $E(r)$ is the radial component of the electric field averaged over a single orbit at radius r. The field free result is obtained by measuring on several ions of different mass and extrapolating to zero mass. In order to do this, it must be assumed that $E(r)$ remains unchanged from ion to ion. Franken has pointed out that the space charge component of $E(r)$ may vary with ion for ions which compose a large fraction of the space charge. If one assumes approximately constant sensitivity of the omegatron with change in mass one can find the ions whose cycling would not disturb the $E(r)$ value significantly simply by comparing the maximum sized signals for the several ions. Principally, from this consideration the ions H^+, D^+, and HD_2^+ were selected to perform the zero mass extrapolation because each was less than 5 percent of the major peaks. HD_2^+ was selected rather than D_3^+ because of the absence of ions of approximately half of its mass in the background gas. There was some concern over neighboring peaks influencing each other and that peaks present at higher harmonics might cause a shift in the center of the line. No evidence was found that it was important under the conditions in which data was taken with a gold omegatron (i.e., drastically reduced trapping field and slightly reduced rf electric field).

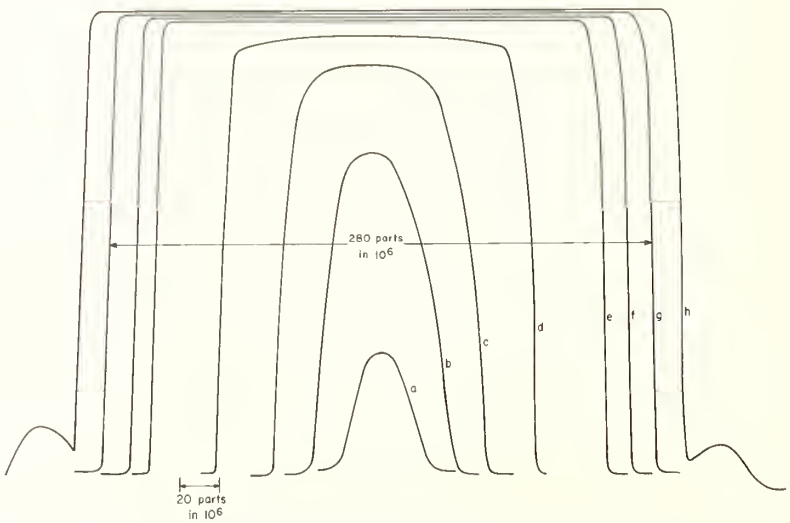


FIGURE 4. *Lineshapes showing variation with change of rf voltage.*
These are for H_2^+ ions. The several lines have been smoothed slightly in drawing to make interpretation easier.

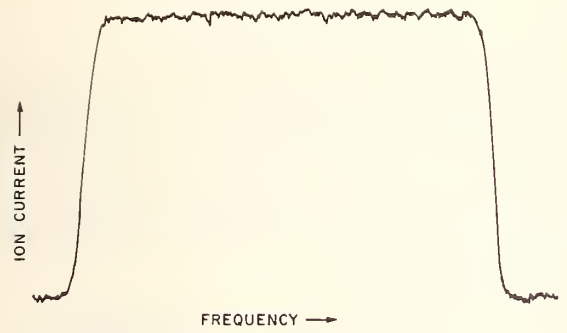


FIGURE 5. *Most rectangular line for protons of a set such as shown in figure 4.*
Line width at the base is about 110 ppm. No smoothing has been done in drafting of this figure. The shift of the center of this line is about 0.5 ppm. between 15% and 85% of maximum.

TABLE 1. *Tabulation of errors*

Source of Error	
Magnetic field inhomogeneity	±4 ppm
Surface potential variations	±3 ppm
Magnetic Contamination	±1 ppm
Electric field penetration from electron collector holes	±1.5 ppm
Rf voltage variation across line	±0.5 ppm
Frequency measurement	±0.1 ppm
Uncertainty in correction from filament current shift	±0.1 ppm
Uncertainty in relativistic mass correction	±1 ppm
Uncertainty in mass ratios	±0.1 ppm
Line asymmetries	±1 ppm
Rf field inhomogeneity	±0.5 ppm
Standard error of data	±1.2 ppm
Total Root Sum Square of random and systematic error	±6 ppm

3.1. Zero Mass Extrapolation Performed

Figure 6 shows this zero mass extrapolation for 9 sets of ions corrected to the same magnetic field. $\gamma_m A_m f_m$ is the observed cyclotron frequency f_m times the atomic mass A_m with the corresponding relativistic correction γ_m. The range of the extrapolated value is about 3 ppm with a standard deviation of 1.2 ppm.

4. Uncertainties

The largest uncertainty in this work is from the magnetic field inhomogeneity which was ±6 ppm over the active volume of the omegatron. Because the actual orbits of the ions are not known the 70 percent confidence interval from this source is given at ±4 ppm. With the presently used crude orbit theory a more homogeneous magnetic field would have been required to reduce this uncertainty, but no high homogeneity magnet was available with a large enough gap to be compatible with the omegatron used. Work was halted when other uncertainties had been reduced substantially below the uncertainties due to field inhomogeneity and to surface potential variations.

4.1. Surface Potential Variations

Shifts from surface potential variations are not necessarily completely corrected for by the extrapolation procedure discussed in 3.1 because the ions of different mass may average over slightly different orbits and see slightly different average electrostatic fields. This could originate from different motions along the magnetic field direction and/or a difference in paths in which ions of different mass move because of orbit drift. A comparison of results of a determination under two quite different surface conditions served as the basis for an estimate of the size of this uncertainty. The first measurement was made after a heavy oxide coating on the omegatron had been partly and asymmetrically removed by a hydrogen discharge. The omegatron was then cleaned and gold plated before a second determination was performed. The difference between the

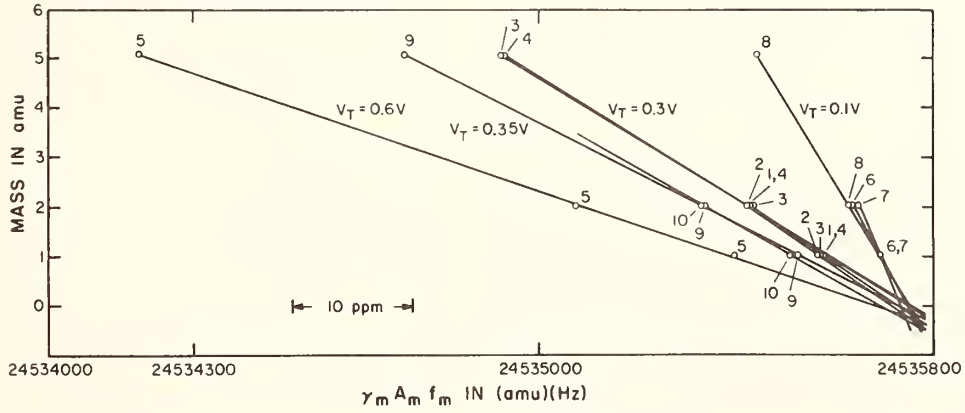


FIGURE 6. *Extrapolation to zero mass for the data taken with the omegatron gold plated.*
Mass one ions were H_1^+, mass two D^+, and mass 5 ions were HD_2^+. The range of the intercepts is 3 ppm. This intercept is the proton cyclotron frequency times the rest mass of the proton. Points of the same run have the same number. V_T is the trapping voltage.

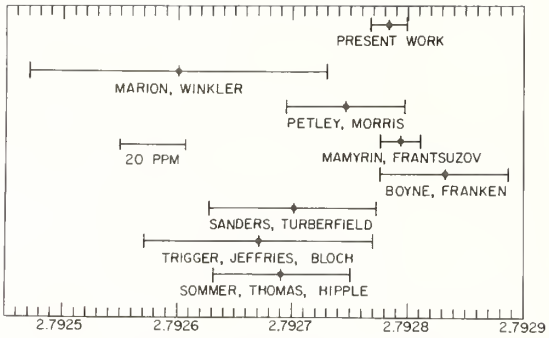


FIGURE 7. *Comparison of results.*
All the error bars are believed to represent 70% confidence intervals.

two was approximately 3 ppm and ±3 ppm is taken as the uncertainty interval from surface potential variation.

4.2. Tabulation of Uncertainties

Table 1 shows a tabulation of all error assignments. The square root of the sum of the squares (RSS) of the random and systematic errors is ±6 ppm. This ±6 ppm represents a 70 percent confidence interval of the experiment.

5. Corrections Made to the Final Result

Corrections have been made to the result for filament current magnetic field and relativistic mass shift. The filament correction was made by taking half of the difference between the results for filament current forward and filament current reversed and is 0.4 ppm. The relativistic correction was significant only for the proton and deuteron, being approximately 3 ppm and 0.8 ppm respectively. Uncertainty in this correction results from the possibility of orbit drift.

6. Final Result

The final result obtained for μ_p'/μ_n in the form uncorrected for diamagnetism of the nuclear resonance sample is 2.792783(16) or ±6 ppm. This represents a 70 percent confidence interval and is the RSS of the systematic and random errors. This result is an average with equal weight of the results of measurements taken with the omegatron before and after gold plating.

7. Comparison of Results

A comparison of results is shown in figure 7 with those determinations previous to 1969. The result reported here is closest to that of Mamyrin and Frantsuzov [4] and is in strong disagreement with the final recommended value for μ_p'/μ_n of Taylor, Parker, and Langenberg [5].

8. Acknowledgments

I am happy to thank Peter L. Bender for his guidance and help in this work. The magnet and space for carrying out the experiment were provided by the NBS Radio Standards Physics Division. The omegatron was constructed by William B. Halliday and reworked by Richard Weppner.

9. References

[1] Sommer, H., Thomas, H. A., and Hipple, Phys. Rev. **82**, 697 (1951).
[2] Robinson, F. N. H., J. Sci. Inst. **34**, 1260 (1959).
[3] Berry, Clifford E., J. Appl. Phys. **25**, 28 (1954) and also Shabad, B. O., Zhurnal Tekhnicheskoi Fiziki **36**, 720 (1966).
[4] Mamyrin, B. A., and Frantsuzov, A. A., *Proceedings of the Third International Conference on Atomic Masses*, edited by R. C. Barber (University of Manitoba Press, Winnipeg, 1968), p. 427.
[5] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

Measurement of μ_p/μ_n

B. W. Petley and K. Morris

Quantum Metrology Division, National Physical Laboratory, Teddington, Middlesex, England

This paper reports recent measurements of μ_p/μ_n with a quadrupole omegatron. The measurements have been made at a flux density of 1.4 tesla and ions of mass 2, 3 and 4 have been used to eliminate the shifts in the cyclotron frequency by electrostatic field gradients. The result of the present series of measurements is that μ_p/μ_n is 2.792828 ± 0.000023 (8.2 ppm) and 2.792754 ± 0.000023 (8.2 ppm) with and without the correction for the diamagnetism of water respectively.

Key words: Atomic constant; magnetic moment; μ_p/μ_n; precise measurement; proton.

1. Introduction

The magnetic moment of the proton is measured as the ratio of the proton spin precession frequency to the proton cyclotron frequency in the same magnetic flux density [1]. The spin precession frequency has been measured using the amplitude bridge method of Thomas and Huntoon (1949) [2] with a cylindrical water sample. The major experimental difficulty arises from the measurement of the proton cyclotron frequency.

Sommer, Thomas, and Hipple [3] devised the omegatron method for measuring the proton cyclotron frequency. The method suffers from the disadvantage that the ion trapping field is obtained by applying potentials to the electrodes that surround the analysis region. If these are plane surfaces the expression for the trapping potential distribution in the analysis region contains significant high order terms and the equations of motion can only be solved by approximate methods [4]. Our quadrupole omegatron [5, 6, 7] represents the simplest omegatron design that both provides ion trapping and has equations of motion that have analytic solutions. This paper discusses the progress that has been made towards the measurement of the ion cyclotron frequency with a quadrupole omegatron.

2. The Quadrupole Omegatron

In our omegatron, figure 1, the electrodes bounding the analysis region are bent to a hyperbolic shape. The trapping potential distribution is made a function of two space co-ordinates only by using a guard disc at either end to simulate an infinitely long device. The guard discs in our 1967 determination [5, 7, 8], were fabricated from glass and the inner surfaces made conducting by means of a NESA coating.

Two of the hyperbolic electrodes were split into four sections and these were at the same potential for direct voltages but were connected through dropper resistors for their radio-frequency voltages. The best relative values for the radio-frequency voltages were found from an electrolytic tank study of the rf field uniformity. The rf field was therefore uniform to better than 10 percent over the region covered by the ions. The leading term in the expression for the non-uniformity was of course the linearly increasing field resulting from the electrode shape. This term may be varied experimentally to study its effects, by varying the potential tappings of the guard plates.

The 1967 determination confirmed that the performance of the quadrupole omegatron led to a unique value for μ_p/μ_n in spite of the presence of asymmetries in the resonance. These asymmetries were a result of the variations of the magnetic field, variations of the residual electrostatic fields over the ion orbit and were not necessarily a consequence of using a quadrupole omegatron.

The variations in the magnetic field have been largely eliminated in the present work by using an electromagnet of greatly improved performance. The variations in the electrostatic fields over the orbit were thought to be due in part to variations of the contact potential difference between the omegatron electrodes and the coated glass end guard plates.

A number of improvements have been made in the omegatron design used for the present series of measurements. The filament leadout wires were brought closer together thereby reducing the magnetic field produced by the heater current. The omegatron was constructed from 80/20 platinum-iridium and made 50 percent longer in order to lessen the proximity of the end plates to the ion analysis region. The end plates were also shielded by a hyperbolic star made of platinum-iridium and shaped pieces were welded to the other electrodes. An electrolytic tank plot showed that the fields at the analysis region were not affected by this modification.

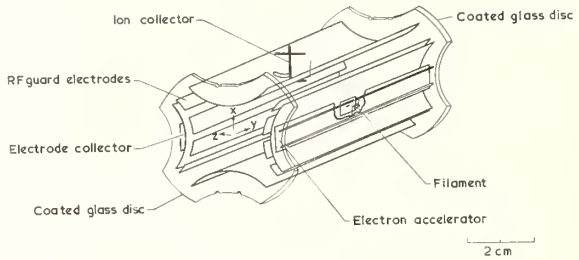


FIGURE 1. *Cut-away view of a quadrupole omegatron (the stars guarding the glass end plates are not shown).*

The electron beam was of a rectangular shape ($\frac{1}{2}\times 1$ mm). The ion collectors were off-set by 4 mm and penetrated 1 mm into the analysis region.

3. The Magnet

A 15-in. electromagnet has been used which gives a flux density of 1.4 tesla in a 5 cm gap. The magnetic flux was stable to within a part in a million over more than 30 min and the uniformity achieved, after switching on and without any cycling, corresponded to about 1.5 ppm over a cube of side 3 cm (the distance between the omegatron rf plates was 2 cm).

The magnet was water cooled and its temperature maintained constant to within ± 0.01 °C.

4. Other Apparatus

As with our previous determination the omegatron was pre-evacuated. The gas pressure could be adjusted by a sputter-ion pump at the beginning of each day. An evaporation ion pump was employed to keep the pressure constant while measurements were being recorded. Typical indicated operating pressures were $\lesssim 40$ μPa (0.3 μtorr) and once stabilized, the pressure remained constant to <1 percent while a set of resonances were being obtained.

Battery operated crystal oscillators were used to generate the 60 MHz spin precession frequency and the corresponding ion cyclotron frequencies. The omegatron power supplies were also battery operated. The frequency counter, electrometer and x–y pen recorder associated with the observation of the cyclotron resonance were mains powered. The phase-sensitive detector (and frequency counter) for the observation of the spin precession frequency and the ionization gauge and sputter pump supply were also mains operated.

5. Method

Each resonance was scanned at a fixed applied radio frequency by varying the magnetic field. A signal derived from the magnetic field sweep was connected to the x-input of the pen-recorder and the collected ion current, represented by the electrometer output, was connected to the y-input. A marker pulse, corresponding to the centre of the proton resonance signal, was superimposed on each trace. Once the x-calibration of the pen-recorder had been established in terms of change in the proton spin-precession frequency, the distance between the marker pulse and the centre of the cyclotron resonance could be established as a change in the spin-precession frequency f_s.

Figure 2 is an example of the type of cyclotron resonance signal recorded. It is apparent that the rectangular curve of the simple omegatron theory is essentially observed; the differences arise as follows: the curvature across the top of the resonance is the result of collision losses in the ion intensity because ions at the centre of the resonance travel a shorter distance than those at the resonance edge (by a factor $2/\pi$). The slope at the resonance edge is a direct result of the finite electron beam size. The base-width of the resonance corresponds to about 360 ppm for the He^+ resonance and would be half of this for the H_2^+ resonance.

Although the asymmetry in the resonance was barely detectable it was decided to make measurements of the centre of the resonance at every 10 percent position from the base to the top of the resonance.

The shifts in the proton cyclotron frequency were eliminated by recording the resonance curves for H_2^+, HD^+, and He^+ ions, taking about 2.5 min to reeord each resonance. Conditions were sufficiently

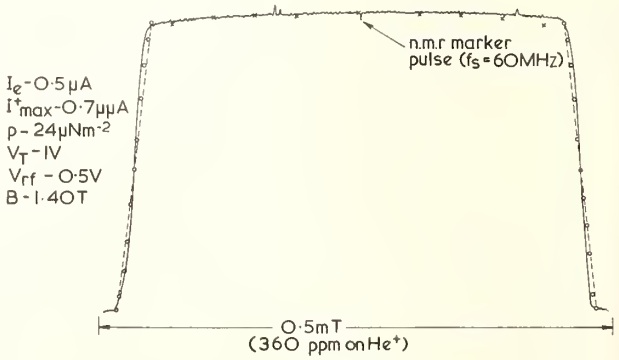


FIGURE 2. *Illustrating the type of cyclotron resonance curve recorded.*

The resonance shape results from (i) collision losses (×—points for shape when the resonant ion path length is 0.1λ) and (ii) finite electron beam size (○—circular electron beam; ---—square electron beam). An actual He^+ recorded resonance is superimposed (solid line).

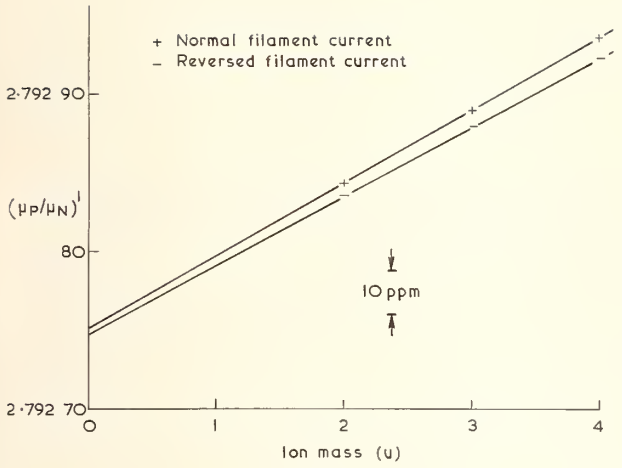


FIGURE 3. *The computed values for* (μ_p'/μ_n) *(without the corrections added) for* H_2^+, HD^+, *and* He^+ *for all of the runs.*
The points shown are for measurements made at the resonance half-height for the two directions of the filament current.

stable for all three resonances to be recorded on the same sheet of paper.

Figure 3 illustrates the method of eliminating the systematic frequency shift, showing the average of the 50 percent height measurements for all of the runs. Runs were recorded for the two directions of the heater current and the effect of the small magnetic field from this current is clearly visible as a difference between the intercepts.

The standard deviation of a single intercept was slightly greater for measurements at the top and bottom of the resonance than for the centre of the resonance, about 5.5 ppm and 3.5 ppm respectively. Sufficient observations were made to reduce the standard deviation of the mean to about 1 ppm.

6. Results

The results of measurements of 42 sets of resonances obtained under different operating conditions of pressure, rf voltage, trapping voltage, electron current and filament current direction are that the uncorrected value of μ_p/μ_n is $2.792\ 746_2 \pm 0.000\ 000\ 71$ and $2.792\ 749_8 \pm 0.000\ 000\ 56$, for the filament current in the negative and positive directions respectively. To these must be added the magnetic field and other local corrections (table 1) and also the systematic errors (table 2). At present the total error, estimated as a standard deviation, corresponds to 8.2 ppm. The corrected result is that the proton magnetic moment, in terms of the nuclear magneton is $2.792\ 828 \pm 0.000\ 023$ (8.2 ppm) with and $2.792\ 754 \pm 0.000\ 023$ (8.2 ppm) without the correction for the diamagnetism of water respectively.

This value is within 3 ppm of our originally published result and it is interesting that we obtained a similar unpublished value with the present omegatron, using our old electromagnet. The flux density for the latter determination was at $0.47T$, compared with the present flux density of $1.4T$. Thus it appears that any unknown systematic errors do not depend markedly on the magnetic flux density. This agreement does not of course exclude systematic errors that are common to both of the latter experiments. Work will continue both on using the omegatron and on evaluating the corrections.

TABLE 1. *Corrections to be added to the measured* μ_p/μ_n values

	ppm
Relativistic mass increase	−1.1
Omegatron 'shielding'	+0.25
Fixed nmr probe to average magnetic field over the ion orbit	+2.15
Shielding of nmr probe	+2.5
Sample shape correction (cylindrical to spherical)	−1.5
	+2.3
Diamagnetic correction (assumed)	+26.0
	+28.3

TABLE 2. *Errors (expressed as standard deviations)*

	ppm	$(ppm)^2$
Estimate of cyclotron resonance centre	4.0	16.0
Possible skewness of resonance	3.0	9.0
Elimination of electrostatic field frequency shifts	6.0	36.0
Relativistic mass increase	0.5	0.25
Elimination of filament current effect	0.2	0.04
Standard deviation of mean intercept	0.4	0.16
Magnetic screening correction	2.0	4.0
Diamagnetic correction	1.0	1.0
Frequency measurement	0.1	0.01
R.S.S.	8.2	66.46

7. Acknowledgment

This work forms part of the research programme of the National Physical Laboratory.

8. References

[1] Alvarez, L. W., and Bloch, F., Phys. Rev. **57**, 111 (1940).
[2] Thomas, H. A., and Huntoon, R. D., Rev. Sci. Instrum. **20**, 516 (1949).
[3] Sommer, H., Thomas, H. A., and Hipple, J. A., Phys. Rev. **82**, 697 (1951).
[4] Petley, B. W., and Morris, K., J. Phys. E **1**, 417 (1968).
[5] Petley, B. W., and Morris, K., Proceedings of the Third International Conference on Atomic Masses, R. C. Barber, Ed. (University of Manitoba Press, 1968), p. 461.
[6] Petley, B. W., and Morris, K., National Physical Laboratory, Quantum Metrology Report No. 7 (1969).
[7] Petley, B. W., and Morris, K., J. Sci. Instrum. **42**, 492 (1965).
[8] Petley, B. W., and Morris, K., Nature **213**, 586 (1967).

A Novel Type Determination of the Magnetic Moment of the Proton in Units of the Nuclear Magneton, μ_p/μ_n

H. Gubler, W. Reichart, M. Roush,* H. Staub, and F. Zamboni

University of Zürich, Switzerland

Previous measurements of μ_p/μ_n were all based on the determination of the cyclotron frequency of protons circulating in a homogeneous magnetic field whose value was obtained through its proton nuclear spin resonance frequency. In the present experiment, the momentum of highly monochromatic protons (energy 1.3 MeV with less than 100 eV spread) is determined in a 180° magnetic spectrometer whose field is measured by its proton nuclear spin resonance frequency. The velocity of the protons is measured with the Los Alamos time of flight spectrometer through which the protons pass before entering the magnetic spectrometer. The results are expected to have an accuracy of at least 20 ppm.

Key words: Magnetic spectrometer; μ_p/μ_n; time of flight analyzer; velocity gauge.

1. Introduction

Marion and Winkler [1] compared in 1967 some values of nuclear reaction calibration energies obtained with a time-of-flight analyzer (velocity gauge) at the University of Maryland with those obtained with an absolute magnetic analyzer at the University of Zürich. The slight differences of the results could be ascribed to a small error in the value of μ_p'/μ_n used in the analysis of the data. However, the evidence was not conclusive, since target effects as for instance surface contamination, Lewis effect, and channeling can amount to several times the observed difference.

In order to avoid errors, measurements with the two instruments were taken simultaneously on one and the same beam of particles with an energy of about 1.3 MeV, after transferring the velocity gauge, built by Seagrave [2] in Los Alamos, to the University of Zürich.

Figure 1 shows the experimental set-up.

A proton beam from our 5.5 MeV electrostatic accelerator is deflected by a 90° magnet and its energy spread reduced to 60 eV/MeV by a fast stabilizer [3]. The particle velocity is determined by the velocity gauge. The value of the magnetic field of the magnetic analyzer in terms of the proton nmr frequency is determined by deflecting the particles of known momentum on a precisely defined orbit.

2. Velocity Gauge

The velocity gauge is a time-of-flight beam analyzer. The ion beam is swept by a (35 MHz) rf-voltage over a small aperture. The particle bunches enter a coaxial cavity and excite electromagnetic oscillations upon passing the two gaps located at the input and the output of the cavity. If the cavity is tuned at about the repetition frequency of the bunches, a pickup antenna in the cavity detects maxima or minima of intensity, depending whether the pulses at the two gaps are in phase or out of phase relative to each other.

The detection of minima as a null method is obviously superior. The normal mode frequencies of the cavity enter only through their effects on the Q-value of the cavity and thereby on the shape of the resonance. The effects of tuning and higher cavity normal frequencies on the position of the minima are entirely negligible. If each gap is symmetric with respect to a plane perpendicular to the beam axis, the effective length is just the distance between these planes of symmetry. We have made some modification in the construction of the ends of the cavity, to be able to measure the length with high precision, and to guarantee good mechanical stability over long time.

A beam collimator reduces the current on the internal part of the cavity (drift-tube) to less than 1/2000 of the current going through the gauge (max. about 2 μA), and correspondingly increases the sharpness of the approximately parabolic minimum resonances.

3. Magnetic Analyzer

The magnetic analyzer built by Winkler and Zych [4] is a magnet of 50 cm radius, whose slits of new design are cut into a 14 mm molybdenum rod. The slit apertures are about 15 μ and they lie in the homogeneous part of the magnetic field. The slit

* University of Maryland, Physics Department, College Park, Maryland 0740.

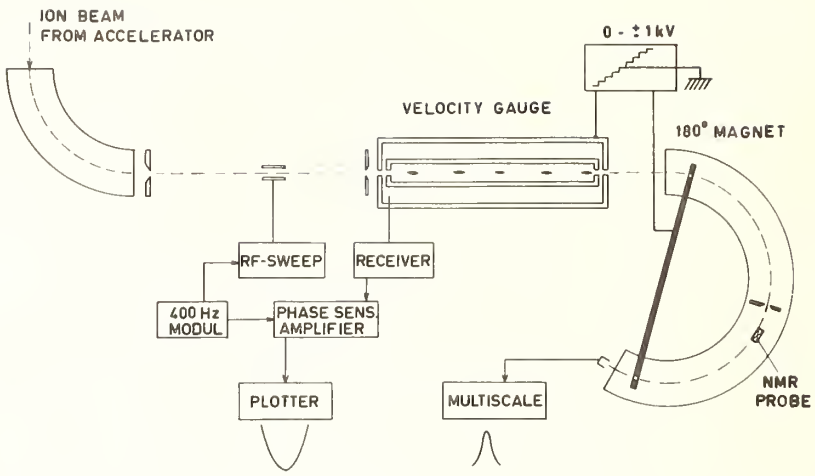


FIGURE 1. *Experimental set-up.*

is temperature controlled, and the temperature gradient between input and output slit is less than 1 °C (thermal expansion coefficient of Mo is 4.9 ppm/°C). The aperture angle is limited by the central slit which is 1.5 mm wide. The field is stabilized by a nmr probe whose signal is locked-in electronically. The long time stability is about 2 ppm. The field along the orbit is corrected with magnetic tape shims positioned in the gap. Its configuration along the orbit of the particles is measured at the beginning and at the end of every complete run by a movable nmr probe, which measures the difference between the nmr frequency of the stabilizing probe and the frequency values along the orbit. From this value the effective field is easily determined (Hartree correction). (Fig. 2.)

$$B = \frac{1}{2}\int_0^{\pi} B(\varphi)\,\sin\varphi d\varphi$$

The Hartree correction at the beginning and the end of a run differed always by less than 5 ppm. The radial field gradient is about 7 ppm/mm.

4. Automatic Recording of Data

The signal picked-up by the antenna in the cavity is amplified by a receiver tuned at the resonance frequency and selected from the background by a phase sensitive amplifier. The particles at the output of the magnetic analyzer are scattered by a Ni foil and counted by a solid-state detector at a fixed angle.

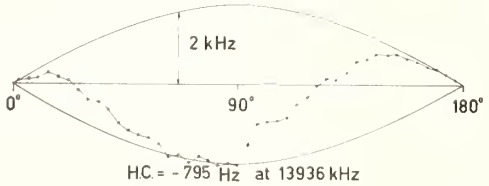


FIGURE 2. *Hartree correction.*

In order to avoid changing the accelerator voltage which always causes small changes in the beam geometry, we have applied a step voltage to both velocity gauge and magnet vacuum chamber. This step voltage triggers the plotter which records the output signal from the phase sensitive amplifier (fig. 3). The pulses from the solid state detector are accumulated in a multichannel analyzer which operates in the multiscale mode and is also triggered by the step voltage device (fig. 4). A single run consisted of a complete set of about 20 simultaneous measurements of minima (nulls) of the velocity gauge and beam profiles of the magnet output. The

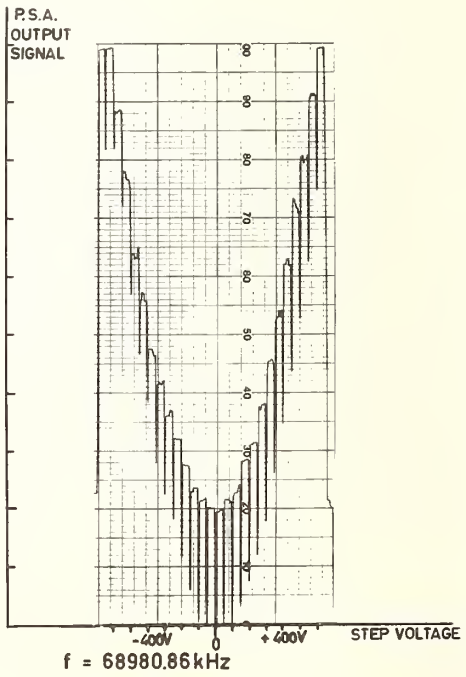


FIGURE 3. *Velocity gauge resonance signal.*
The spikes are step voltage marks.

position of the minimum is determined by least square fit. The midpoint at FWHM of the beam profile was chosen as the mean velocity. The statistical error of a set of measurements of a run as less than 3 ppm.

5. Analysis of Data

The condition for minimum signal of the velocity gauge is

$$v=2fL/(2N-1) \qquad N=1,\ 2,\ 3\cdots \text{ Mode number},$$

where f is the frequency of modulation of the beam, L is the distance between the planes of symmetry of the gaps and v is the velocity of the particles.

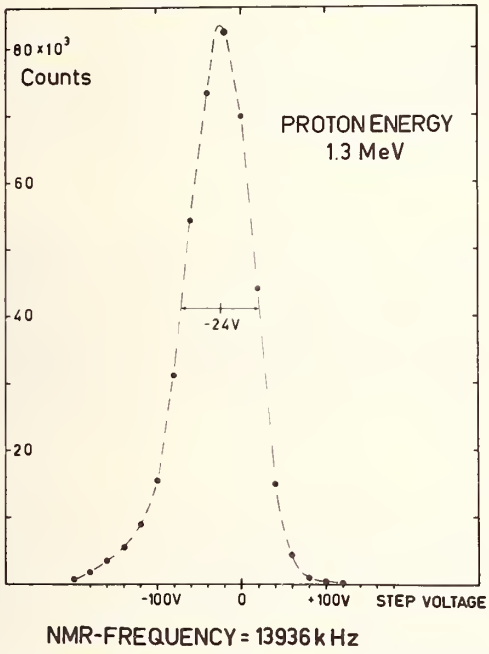


FIGURE 4. *Beam profile taken by the magnetic analyzer.*

The relation between velocity and field in the magnet is given by

$$B=m_p v/(1-\beta^2)^{1/2}\cdot(c/eR)$$

and

$$B=\omega_{\mathrm{nmr}}/\gamma=\omega_{\mathrm{nmr}}(\hbar/2\mu_p),$$

where m_p is the proton rest mass, ω_{nmr} the proton nmr frequency corresponding to the field B, and γ the gyromagnetic ratio of the proton.

It follows:

$$\mu_p'/\mu_n=\pi\tfrac{1}{2}(2N-1)\,(D/L)\,(\nu_0/f)\,(1-\beta^2)^{1/2}$$

$$\nu_0=\omega_{\mathrm{nmr}}/2\pi,$$

where D is the slits separation, ν_0 is the nmr frequency for protons in a spherical sample of water, and:

$$\beta=2fL/(2N-1)c.$$

TABLE 1

Error	ppm
—Fluctuations of position of minima and beam profile centers	3
—Hartree correction	5
—Uncertainty in the position of particle orbit	6
—Frequency	2
—Geometrical length L	2
—Slit separations D	2
—Slit apertures	4
—Slit temperatures	2.5
—Drift tube temperature	1
—Vacuum	1.5
—Drift tube alignment	1.5
Total (RSS)	11 ppm

Introducing the Hartree correction and the energy changes due to the step voltages we get:

$$\mu_p'/\mu_n=\pi\tfrac{1}{2}(2N-1)\,(D/L)\,(\nu_0/f)\,(1-\beta^2)^{1/2}$$
$$\times\{1+(\Delta\nu_H/\nu_0)+\epsilon[(1-\beta^2)^{1/2}/\beta^2]\},$$

where $\Delta\nu_H$ is the Hartree frequency correction and

$$\epsilon=(V_M-V_{VG})\,(e/m_pc^2),$$

with

V_M = magnet step voltage, and
V_{VG} = velocity gauge voltage.

Note that the lengths D and L and the frequency ν_0 and f need not to be measured absolutely, since only their ratio enters.

6. First Results

The new construction of the cavity end plates and of the magnet slit system made it possible to measure the two lengths with the necessary accuracy at the "Eidgenössische Amt für Mass and Gewicht" in Bern. The error is about 2 ppm. The sources of uncertainty and their standard deviations are given in the following table 1.

A small correction of few ppm has to be applied for the nonspherical shape of the nmr probe and the effect of the 0.05 molar paramagnetic ions added in the probe for shortening of the relaxation time.

Up to now we have carried out 5 runs. Their results are listed in table 2:

TABLE 2

Present Results	
Run 1	$\mu_p'/\mu_n=2.7927\ 58$
Run 2	$=2.7927\ 90$
Run 3	$=2.7927\ 57$
Run 4	$=2.7927\ 70$
Run 5	$=2.7927\ 80$
Mean value	$\mu_p'/\mu_n=2.7927\ 71\pm0.0000\ 31$ (11 ppm)

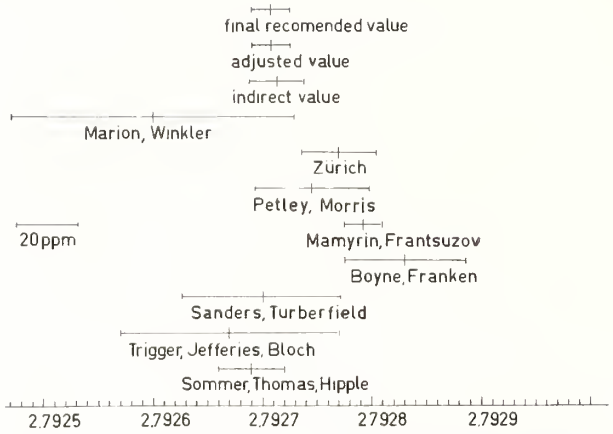


FIGURE 5. *Comparison of our present result with the previous ones.*

7. Discussion

Figure 5 shows our present result, compared with the previous ones. This result is still subject to the following possible systematic errors, which will be carefully investigated in the near future. Errors could arise from the following:

(1) The beam orbit differs slightly from that measured with the movable nmr probe.
(2) A small asymmetry in the cavity gaps, which changes the value of the effective cavity length.
(3) An error in the measurements of L and D.

We believe that the combined effect of these three errors should not change the above result by more than 15 ppm. In particular the following investigation will be carried out:

(1) Repetition of measurements with different particle energies which encompasses a change of the magnetic field configuration.
(2) Measurements with different geometry of the gap defining inserts.
(3) Repeated disassembly of the cavity and remeasurements of its lengths.

The final results will be published in a more detailed paper.

8. Acknowledgments

This work has been supported financially through the Schweizerische Nationalfonds zur Förderung der wissenschaftlichen Forschung. We gratefully acknowledge the cooperation of the Eidgenössische Amt für Mass und Gewicht in Bern.

9. References

[1] Marion, J. B., and Winkler, H., Phys. Rev. **156**, 1062 (1967).
[2] Seagrave, J. D., Brolley, J. E., and Beery, J. G., Rev. Sci. Instr. **35**, 1290 (1964).
[3] Bloch, R., Pixley, R. E., Reichart, W., and Zamboni, F., Nucl. Instr. Meth. **59**, 325 (1968).
[4] Winkler, H., and Zych, W., Helv. Phys. Acta **34**, 450 (1961).

DISCUSSION

V. W. COHEN: Neither of the previous speakers mentioned the nature of the molecular correction. Presumably they measure the proton resonance in water. But how precisely are the molecular corrections known at this time?

F. ZAMBONI: We always refer to measurements made with a spherical water probe according to the definition of μ_p'. Occasionally we have to apply a correction for the cylindrical shape of the probe and the contribution of the paramagnetic ions, but the results are always given for a spherical probe with pure water.

A New Technique for Measuring the Cyclotron Frequency of Ions to High Precision

James L. Luxon and Arthur Rich

The Harrison M. Randall Laboratory of Physics, The University of Michigan, Ann Arbor, Mich. 48104

We have developed a new technique for measuring ion cyclotron resonance. This technique may be applied to an accurate measurement of the magnetic moment of the proton in nuclear magnetons μ_p/μ_n as well as to ion mass ratios. The errors in applying this method should be different in character from those which occur in previous cyclotron resonance experiments.

Key words: Ion cyclotron resonance; ion mass ratios; magnetic moment of the proton in nuclear magnetons.

We have recently developed a new technique for measuring ion cyclotron resonance. It is being applied to a precision measurement of the magnetic moment of the proton in nuclear magnetons μ_p/μ_n. Our technique should allow us to measure ion cyclotron frequencies to an accuracy of a few parts per million without any significant contributions to the error from electric field shifts, relativistic mass corrections, or the natural line width of the resonance. The major source of error is now anticipated to be due to the inhomogeneities in the magnetic field. These inhomogeneities are deliberately introduced so that the field will confine the ions.

We will first briefly describe the technique. A "bunch" of ions is trapped in a shallow magnetic mirror well [1]. In the well the trapped ions execute both cyclotron motion and a slow oscillatory drift back and forth between the mirror points. While the ions are trapped, an rf field is applied across their orbits. When the rf frequency is equal to the time averaged cyclotron frequency of an ion, the ion undergoes a small continuous energy change. This energy change implies a change in radius (about 10 percent) which is used to detect resonance.

There are two primary advantages in this method as compared to the methods used in previous cyclotron resonance measurements [2]. First, trapping allows us to keep the ions in the rf region for a large number of cyclotron revolutions. This allows the natural linewidth of the resonance to be made very small. Second, the electric field shift of the cyclotron frequency is minimized by using magnetic trapping, isolating the experimental region from the electric fields of the source and detector, and eliminating the electric fields due to space charge.

This technique may be applied to a measurement of the magnetic moment of the proton in nuclear magnetons μ_p/μ_n by using the relation $\mu_p/\mu_n=f_s/f_c$, where f_s, the spin precession frequency of the proton, is obtained by averaging an NMR map over the experimental region and $f_c=eB/2\pi\gamma M_p c$ is the cyclotron frequency of the proton [3]. We should be able to measure both f_s and f_c to a few parts per million.

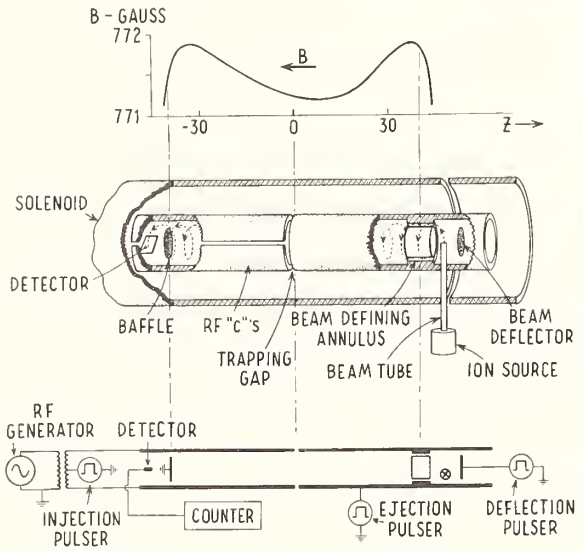


FIGURE 1. *The experimental apparatus with the approximate magnetic field and the basic electronic connections.*

The apparatus can also be used to determine ion mass ratios. In particular, mass doublets such as H_2^+–D^+ which have nearly the same mass move through the apparatus with almost identical orbits. Therefore, shifts in the cyclotron frequency due to systematic effects cancel. This allows a determina-

tion of the relative separation of the two lines to a higher accuracy than we can measure an absolute cyclotron frequency.

The essentials of the experiment are shown in figure 1. A large solenoid produces a magnetic field of about 800 gauss. Additional coils shape the field to produce a shallow magnetic mirror well. The field is regulated to a part per million using NMR techniques. Within the solenoid is a vacuum system which maintains a pressure of 2×10^{-9} mm Hg. The vacuum system contains a set of cylindrical electrodes which surround the experimental region. These can be visualized as one long cylinder cut in half across the axis near the center of the well. One of the halves is then split lengthwise and connected to an rf source. To the right of the well is the beam tube. This tube directs the ions into the center of the solenoid from an external Ortec ion source. After the ions leave the beam tube they move through an annulus which defines their radius to be $r=4.15\pm 0.15$ cm, and into the well region where they are trapped. When the ions are ejected from the trap they move to the left past another baffle and on to the detector. The detector is a Bendix type magnetic strip multiplier which utilizes the magnetic field of the solenoid and operates in the particle counting mode. It has a variable aperature and can be adjusted over a wide range of radial positions.

The effect of the magnetic field configuration as an axial potential well can be easily shown. First, we define $v_z(z)$ as the component of the particle's velocity parallel to the magnetic field, and $v_\perp(z)$ as the component of the velocity in the transverse plane. Then, since the speed of the particle is constant in a purely magnetic field, we can write

$$v_z^2(z)+v_\perp^2(z)=v_z^2(0)+v_\perp^2(0). \tag{1}$$

If we neglect relativistic effects and define $T_\perp(z)=\frac{1}{2}Mv_\perp^2(z)$ and $T_z(z)=\frac{1}{2}Mv_z^2(z)$, this becomes

$$T_z(z)+T_\perp(z)=T_z(0)+T_\perp(0). \tag{2}$$

Since $\mu=T_\perp(z)/B_z(z)$ is to good approximation an adiabatic invariant in a region of slowly varying magnetic field [4], we can write

$$\mu=T_\perp(z)/B_z(z)=T_\perp(0)/B_z(0). \tag{3}$$

Combining these two equations we get

$$T_z(z)-T_z(0)$$
$$=\Delta T_z=[-T_\perp(0)/B_z(0)](B_z(z)-B_z(0)). \tag{4}$$

This equation explicitly displays the nature of our trap as an axial potential well. Thus, while the ion undergoes cyclotron motion, the orbit center oscillates back and forth in an axial well. The period of this drift motion is 250 μs. Since we used ions with energies of 500 eV in the well shown in figure 1, ΔT is at most 0.3 eV. Finally we note that in the experimental region $B_z(z)$ differs from the total field $B(z)$ by less than a part per million, so that $B_z(z)$ can be replaced by $B(z)$.

In practice the beam system is allowed to run continuously but the ions are given a pitch away from the trap as they leave the beam tube so that they never enter the trapping region. To trap, a positive pulse is applied to the disk, switching the beam toward the well. Simultaneously, the C's electrically connected as a cylinder are raised to a positive voltage, thus removing enough axial energy from the ions as they cross the gap so that they are reflected when they reach the south side of the well. If the voltage on the C's is now removed while these ions are inside, a bunch of ions remains in the well. The disk voltage simultaneously goes back to zero and the beam is again switched away from the well. After the ions are held in the well for a fixed time, they are ejected by raising the north cylinder to a voltage greater than $\Delta T_z(\max)/e$. This gives the ions inside enough axial energy as they cross the gap to leave the well. Using this method we have trapped ions for up to 0.1 s.

The rf process has many similarities to both the omegatron and the cyclotron. While a bunch of ions is confined in the well, an rf field is applied across their orbit by applying an rf voltage given by $V=V_0\cos(2\pi ft+\phi)$ to the C's. The equations of motion have been solved for an ion in this rf field combined with a constant magnetic field. As expected, the solution shows a resonant change in the transverse energy $T_\perp$ of the ion which we shall denote $\delta T_\perp$. This energy change can be written

$$\delta T_\perp=\sum_{n=0}^{\infty}\left[\frac{eV_0}{\pi R}v_0r^{2n}\right.$$
$$\times\frac{\sin[\phi-(2n+1)\theta_0]-\sin\{2\pi[f-(2n+1)f_c]+\phi-(2n+1)\theta_0\}}{2\pi[f-(2n+1)f_c]}$$
$$\left.+\frac{2e^2V_0^2}{\pi^2R^2M_p}r^{4n}\frac{\sin^2\{2\pi[f-(2n+1)f_c]t/2\}}{4\pi^2[f-(2n+1)f_c]^2}\right], \tag{5}$$

where:
$n=0, 1, 2, \cdots$,
V_0 is the peak rf voltage,
v_0 is the initial velocity of the ion,
r is the mean orbital radius divided by the radius of the C's, R,
f is the applied rf frequency,
f_c is the cyclotron frequency,
ϕ is the initial rf phase, and
θ_0 is the initial phase of the particle.

If we let $f=(2n+1)f_c$, this becomes

$$\lim_{f\to(2n+1)f_c}\delta T_\perp=-(v_0eV_0/\pi R)r^{2n}t\cos[\phi-(2n+1)\theta_0]$$
$$+(e^2V_0^2/2\pi^2R^2M_p)r^{4n}t^2. \tag{6}$$

The first term is a phase-dependent energy change which is linear in time. The second term implies an energy gain for all ions and is quadratic in time. If we take $\cos(\phi-(2n+1)\theta_0)=1$, these two terms have the same magnitude when they are separately equal to $4T_\perp$ (initial). Thus, for small energy changes the

first term dominates. In particular, those ions with phase near $\phi-(2n+1)\theta_0=\pi$ will continuously gain energy and spiral to a larger radius, and those ions with phase near $\phi-(2n+1)\theta_0\doteq 0$ will continuously lose energy and spiral to a smaller radius. While these ions change radius, they remain in the magnetic well and must be ejected to be detected. In the process of changing radius, they do experience a small shift in their time-average magnetic field which can be easily calculated. However, the shifts for ions gaining energy and ions losing energy have the opposite sign so that these shifts can be experimentally determined by observing both cases and then interpolating to the initial energy of the ion.

We observe resonance by first establishing trapping with the detector placed at the initial beam radius. The detector is then moved to a radius which is approximately 10 percent smaller or larger than the initial radius (20 percent energy change)

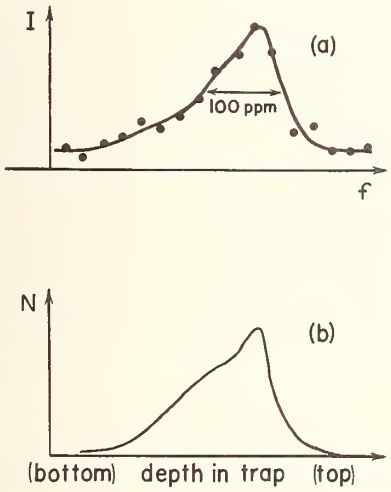


FIGURE 2. *(a) A typical rf resonance line. (b) The population distribution used to obtain that line.*

where we observe only background noise on ejection with the rf off. The rf source is then turned on and set to the voltage indicated by the above equation for $\delta T_\perp$ for the specific trapping time used. When the number of counts obtained on ejection is plotted as a function of the rf frequency, resonance is observed. Resonance has also been observed at the third harmonic of the cyclotron frequency.

The lineshape is determined by two effects. The first is the natural line width of the rf process for a single proton. The first term of eq (5) has a width $(\Delta f/f)_{\text{natural}}=0.60/(2n+1)N$, where Δf is the full width of the line at half maximum intensity and N is the total number of ion revolutions in the rf region. Presently we are trapping protons for 9000 μs with a cyclotron frequency of 1.2 Mhz which gives a natural line width of about 50 ppm (fig. 2). This may be easily improved by an order of magnitude by using longer trapping times and odd harmonics of the cyclotron frequency.

The second contribution to the lineshape is due to the distribution of the ions in the well. Ions at different levels in the well experience different time average magnetic fields

$$[B(z)]=\tau^{-1}\int_0^\tau B(z)\,dt.$$

If, for example, we consider a parabolic well such as the one shown in figure 3 an ion at the bottom experiences $[B(z)]=B_0$ and produces a contribution to the rf spectrum at $f(B_0)$, whereas an ion ΔB higher in the well experiences $[B(z)]=B_0+\Delta B/2$ and produces a contribution to the spectrum at $f(B_0+\Delta B/2)$. In this manner the depth of the well produces a contribution to the linewidth $(\Delta f/f)_{\text{trap}}=1/2(\Delta B_{\max}/B_0)$. For our present trap, $(\Delta f/f)_{\text{trap}}$ can be as much as 300 ppm. Since $(\Delta f/f)_{\text{trap}}$ is generally going to be considerably larger than $(\Delta f/f)_{\text{natural}}$, the shape of the observed line is governed by the distribution of ions at the various levels in the well. However, we have found that we

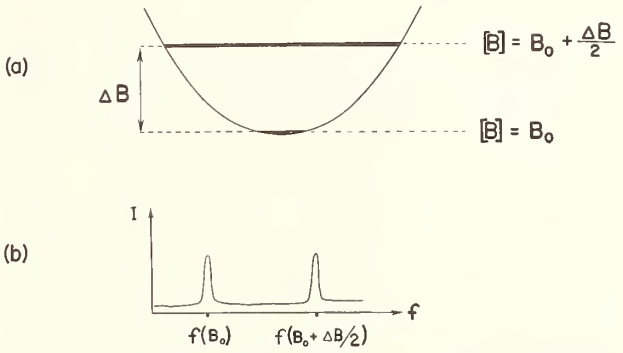


FIGURE 3. *(a) The time-average magnetic field for different levels in an idealized parabolic well. (b) The contributions of these levels to the rf line.*

can control this distribution so that only a fraction of the well is filled. Furthermore, we have developed precise techniques for determining the population of the well [1]. We have found excellent agreement and reproducibility between the population distribution and the shape of the resonance obtained from it as is shown in figure 2. Thus far we have produced population distributions which contribute less than 100 ppm to the linewidth. By decreasing the well depth and improving our techniques for shaping the population distribution, we should be able to reduce this effect considerably.

A second method may also be employed to obtain f_c. This method consists essentially of inverting the order of the steps taken in the procedure described above. We first make a sufficiently accurate determination of the resonance line so that the rf frequency can be set to a value somewhere within the linewidth. As before, the detector remains at a radius 10 percent larger or smaller than the initial radius. On ejection, only those ions which have undergone a resonant energy change reach the detector. These ions must have a time-average cyclotron frequency which is within a natural

linewidth of the rf frequency. The population distribution of this portion of the well may be determined in the same manner as in the previous method. However, it will be much narrower, characterizing only that level in the trap which corresponds to the fixed rf frequency. To obtain μ_p/μ_n we must simply assign the time-average spin precession frequency corresponding to this level in the well. This technique should be more direct and somewhat more accurate than the first technique.

The possibility of a shift in the cyclotron frequency due to electric fields cannot be ignored. If the ion experiences a time-average radial electric field $[E_r]$, the observed frequency f' is given to sufficient approximation for our purposes by the expression $f'=[f_c][1-(c/v)(E_r/B)]$. For the parameters used in our experiment this shift is about 4 ppm per mV/cm. To minimize electric fields, the rf region is shielded from the dc fields of the injection and detection systems by baffles. In addition, by using particle-counting techniques we are able to keep the number of ions in the trap small enough so that electric fields due to space charge are not a problem. Thus, any electric field present would have to be due to potential variations over the surface of the cylinders. This effect is limited by not placing any materials within the beam radius, keeping the ions more than a centimeter from the surface of the cylinders, and coating all of the surfaces with Aquadag collodial graphite. It should be noted that the absence of any materials inside of the orbit along with averaging over the orbit greatly reduces the contribution of any local surface potential variations to $[E_r]$. From previous work done here [1, 5] and elsewhere [6], we would expect the time-averaged radial electric field to be a few mV/cm or less, corresponding to a shift of 10 ppm or less in the cyclotron frequency.

As the experiment continues we plan to correct for this electric field dependence both by measuring f' at several values of magnetic field and by using ions of different mass in a constant magnetic field. If we recall that for cyclotron motion, $r=Mv/eB$, the expression for f' becomes

$$f'=[f_c]\left(1-\frac{Mc^2[E_r]}{er[B]^2}\right). \tag{7}$$

To investigate $[E_r]$, it is desirable to hold r constant, since $[E_r]$ is in general a function of r. However, we can measure f' for ions of different mass while adjusting their energy so that r is constant. If we plot $f'(M/M_p)$ as a function of the ion mass M, we can extrapolate this to $M=0$ to get $[f_c]$. The experiment can also be run at several values of magnetic field while the ion energy is adjusted to hold r constant. If f'/f_s is plotted as a function of $1/B^2$, and this plot is extrapolated to $1/B^2=0$, we again get $[f_c]/f_s$. Using a combination of both of these methods, we should be able to determine the electric field shift well enough that it will not make a significant contribution to our final error.

To measure the magnetic moment of the proton in nuclear magnetons, we also need to measure the proton spin precession frequency f_s. To do this, we map the magnetic field inside of the apparatus at the cyclotron radius using standard NMR techniques on a doped water sample. These measurements are first averaged around the orbit for a given axial position. The values for each axial position are then used to obtain the time-average spin frequency $[f_s]$ experienced by particles at each level in the trap. This frequency is given by the expression

$$[f_s]=\tau^{-1}\int_0^\tau f_s(z)\,dt=\int_{z_1}^{z_2} f_s(z)\,\frac{dz}{v_z}\Big/\int_{z_1}^{z_2}\frac{dz}{v_z}, \tag{8}$$

where τ is the period of the well and z_1 and z_2 are the endpoints. The axial velocity v_z is derived from the axial potential energy relations for the well eq (4) expressed in terms of f_s,

$$\begin{aligned} v_z&=[2T_z/M_p]^{1/2}\\ &=\{[2T_\perp/M_p f_s(0)](f_s(0)-f_s(z))\}^{1/2}. \end{aligned} \tag{9}$$

Using these techniques which are well understood from the positron and electron g-factor experiments performed here [5], we can determine the spin precession frequency to a few parts per million.

In conclusion, we hope to be able to produce accurate measurements of ion cyclotron frequencies without any large systematic corrections. In the case of the proton this should lead to a new determination of μ_p/μ_n. Ion mass doublets should be measurable to sufficient accuracy to provide a useful systematic check of the established techniques [7]. These measurements will be made using a method markedly different from the methods employed in previous work.

Acknowledgments

We thank D. Hartman, D. Pope, B. Wood, and T. Tesar for their help in constructing and operating the apparatus. Personnel of the Bendix Corporation, particularly J. Carrico, M. Johnson, and T. Foote, provided valuable assistance in the modifications of the Bendix strip multiplier for use at our magnetic field strength. This project has been generously supported by the United States Atomic Energy Commission.

References

[1] Our method for the trapping of charged particles in magnetic mirror wells is discussed most recently in Wesley, J. C., and Rich, A., Phys. Rev. Letters **24**, 1320 (1970).

[2] A thorough discussion of these techniques as applied to measurements of μ_p/μ_n may be found in Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[3] Alvarez, L. W., and Block, F., Phys. Rev. **57**, 111 (1940).

[4] S. Chandrasekhar, The Plasma in a Magnetic Field, Ed. R. K. M. Landshof (Stanford Univ. Press, Palo Alto, Calif.).

[5] Wilkinson, D. T., and Crane, H. R., Phys. Rev. **130**, 852 (1963).

[6] Parker, J. H., and Warren, R. W., Rev. Sci. Instr. **33**, 948 (1962).

[7] *Proc. Third Intern. Conf. on Atomic Masses and Related Constants* (Univ. of Manitoba, Winnipeg, Canada, 1968).

DISCUSSION

B. N. Taylor: As we pointed out in our 1969 adjustment in the Reviews of Modern Physics, the major problem facing us at that time was the discrepancies among the various μ_p/μ_n data. I think it is clear from the numbers which were just presented that we probably made the wrong choice and that the true value of μ_p/μ_n is closer to the higher values rather than the lower values, as exemplified by the Sommer, Thomas, and Hipple measurement. This points the finger of suspicion at the faraday as well, and right now it looks like the faraday may be incorrect by something like 25 to 30 parts per million, or four times the standard deviation of the experimental value. Another measurement of the faraday is certainly in order and I am glad to see that a couple of people are trying to settle that problem.

Comments on the Sommer, Thomas, and Hipple Omegatron Measurement of μ_p/μ_n

D. O. Fystrom

Wisconsin State University, LaCrosse, Wisconsin 54601

and

B. W. Petley

National Physical Laboratory, Teddington, Middlesex, England

and

B. N. Taylor*

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

The original data books pertaining to the Sommer, Thomas, and Hipple omegatron measurement of μ_p/μ_n have recently been examined, and new information concerning the uncertainty to be assigned the experiment has been found.

Key words: Fundamental constants; least squares adjustment; μ_p/μ_n; omegatron.

1. Introduction

Since it was completed in 1950, the omegatron measurement of the magnetic moment of the proton in nuclear magnetons (μ_p/μ_n) by Sommer, Thomas, and Hipple [1] has been used as an input datum in all least squares adjustments of the constants. In the most recent adjustment, that carried out in 1969 by Taylor, Parker, and Langenberg [2], the Sommer et al. measurement was not only used as an input datum, but it also played an important role in the decision to expurgate the higher values of μ_p/μ_n obtained by Boyne and Franken and Mamyrin and Frantsuzov. The latter two measurements agreed with each other, but significantly exceeded the Sommer et al. result. Clearly, a substantial increase in the assumed uncertainty in the Sommer et al. measurement might have led to a different conclusion concerning expurgation of these higher values from the Taylor et al. adjustment. Some question about the interpretation of the error given by Sommer et al. did in fact come to light as the Taylor et al. paper went to press and was duly pointed out in their notes added in proof section. The new information came from an examination of the original laboratory data books pertaining to the Sommer et al. experiment and encompassed the following:

(1) The magnetic shielding effect of the omegatron; (2) the proton sample used; (3) the effect of the filament current; (4) the influence of grounding or not grounding the guard rings for rf; and (5) the magnitude and effect of the inhomogeneity of the magnetic field over the active volume of the omegatron. The purpose of this paper is to briefly summarize the information contained in the notebooks as it applies to these areas so that future workers in the fundamental constants field may have it available for their own use and interpretation. Additionally, we give some of our own personal views concerning this experiment.

2. Shielding

During the course of omegatron construction, Sommer et al. [1] checked the magnetic properties of the major components using nuclear magnetic resonance (NMR) to determine the change in magnetic field caused by these components (mainly the guard ring, filament, and electron collector assemblies). It was found that when the filament and electron collector assemblies were held together and placed normal to the field and close to the NMR probe, the magnetic field at the probe was about one

* Work carried out in part while with RCA Laboratories, Princeton, New Jersey.

part in 50,000 stronger than the field measured in the absence of these parts. As a result, Sommer, in his Ph.D. thesis [3], applied a correction to his measured value of μ_p/μ_n of plus 20±4 ppm.

On the other hand, a much smaller correction (plus 2.6 ppm) was applied to the data reported in the final Physical Review paper [1], although not explicitly mentioned in that paper. This correction was obtained from a subsequent experiment on the completed omegatron after the final 45 data runs had been made. It was performed in the following way. Sommer et al. removed the rf electrode, inserted an NMR oil probe inside, reassembled the omegatron, and compared the field at which the probe's precession frequency was equal to that of an adjusted NMR water probe which later replaced the omegatron in the magnet gap. No difference in field for the same frequency was observed. Since it had been separately established that in a given magnetic field, the NMR frequency for the adjusted water sample exceeded that for the oil sample by 2.6 ppm, it was concluded that the omegatron increased the field inside itself by 2.6 ppm. (As a correction to μ_p/μ_n, it contributed plus 2.6 ppm.)

What may be said concerning this apparent difference in shielding,[1] especially since discussions with two of the participants [4] have revealed that no modifications to the omegatron were made between the two tests[2] and that there was only one omegatron in use? We believe that since the double NMR probe test effectively simulated the way the data were taken during the actual experiment, while the first test did not, it is most reasonable to suppose that the 2.6 ppm correction is indeed valid. (See also section 7.) But of course, this cannot be known for certain. A plus 20 ppm correction to the Sommer et al. measurement would in fact bring it into better agreement with the higher values of μ_p/μ_n. (See reference [2] as well as Fystrom and also Petley and Morris, [6] and these Proceedings.)

3. Proton Sample

Although there are no real questions concerning the proton sample used in the Sommer et al. experiment, some additional useful information on the subject has become available from the notebooks.

The value of μ_p/μ_n reported by Sommer et al. [1] was for their "standard oil" NMR probe, although the 45 final data runs were actually obtained using the adjusted water NMR probe, not the standard oil sample. To convert μ_p/μ_n as obtained using the water probe to what would have been obtained had the oil probe been used, a minus 2.6 ppm correction is needed (see section 2). But this is of opposite sign and identically equal to the shielding correction (section 2), so that the two cancel. Thus, the reported value does in fact apply to the standard oil sample but with no shielding correction to be applied.

We should perhaps also remind the reader here that to convert the Sommer et al. result to the form useful for a least squares adjustment, i.e., as determined for protons in a spherical sample of pure water, use must be made of the work of Thomas [5]. He found that for a given magnetic field, the NMR precession frequency of the standard oil sample was 1.9 ppm less than the precession frequency of a spherical water sample. As a correction to the Sommer et al. value of μ_p/μ_n, it is plus 1.9 ppm. (See also [2].)

4. Filament Current Polarity

The tungsten filament used for generating the electron beam in the omegatron was heated with direct current. In two early tests to investigate the effect of this current (actually, the magnetic field due to the current), it was observed that the cyclotron resonance line centers shifted to lower frequency by 8 and 12 ppm, respectively, when the polarity of the filament current was reversed from the usual direction. This would imply that about a plus 5 ppm correction should be applied to the value of μ_p/μ_n. Indeed, such a correction was applied by Sommer in his Ph.D. thesis [3]. However, it was not applied by Sommer et al. to the final data reported in the Physical Review article [1]. This is surprising in view of the results of some special runs carried out after the final 45 data runs were completed.[3] These were made with the same omegatron for the express purpose of investigating the effect of the filament current and grounding or not grounding the guard rings for rf. The results in terms of relative values of μ_p/μ_n are summarized in table 1. The mean of the first two items implies a correction to μ_p/μ_n due to the filament current of plus 7.3 ppm with an uncertainty of about 4 ppm. This is quite consistent with the two earlier tests.

On the other hand, during the course of the final 45 data runs themselves, Sommer et al. made two runs with the filament current reversed and found shifts in the line centers which would imply corrections to μ_p/μ_n of 0 and plus 14 ppm respectively. This is not particularly consistent with the other results. Conversations with two of the participants [4] have revealed that it was probably for this reason that the filament current polarity effect was treated as an error in the final paper rather than as a correction. But upon reexamination of the data, the consensus now is that it might reasonably be applied as a correction. Thus, on the basis of the data in table 1, we will assume that the Sommer et al. value of μ_p/μ_n should be increased by 7.3±4.1 ppm.[4]

---

[1] It might conceivably be related to the difference between a shielding correction and a substitution correction; it is now known that peculiar effects can occur when metal is moved within the gap of an electromagnet.

[2] Other than the possibility that a small piece of nickel may have been used to weld more readily the tungsten filament to the tungstem filament leads during the construction of the omegatron, and the nickel may have been removed prior to the final data runs [4].

[3] The 45 final data runs were carried out from July 19 to August 15, 1950. The special runs were carried out from November 24 to December 11, 1950.

[4] The 7.3 ppm correction has been obtained by taking a simple average of the first two items, and the 4.1 ppm error by dividing the 5.8 ppm pooled standard deviation (all four items) by $\sqrt{2}$.

TABLE 1. *Comparisons of relative values of μ_p/μ_n as measured by Sommer et al. under different omegatron operating conditions.*

The uncertainties quoted are the root-sum-square uncertainties derived from the statistical standard deviations of the means of the runs carried out for the indicated conditions.

Filament current *reversed*, with guard rings *grounded*,	exceeds	Filament current *normal*, with guard rings *grounded*,	by	16.3±5.1 ppm.
Filament current *reversed*, with guard rings *ungrounded*,	exceeds	Filament current *normal*, with guard rings *ungrounded*,	by	12.7±6.5 ppm.
Filament current *normal*, with guard rings *grounded*,	exceeds	Filament current *normal*, with guard rings *ungrounded*,	by	10.8±5.2 ppm.
Filament current *reversed*, with guard rings *grounded*,	exceeds	Filament current *reversed*, with guard rings *ungrounded*,	by	15.5±6.4 ppm.

5. RF Grounded Guard Rings

The final 45 runs were taken with the omegatron's guard rings grounded for rf because this condition gave the least assymmetric cyclotron resonance lines and apparently the most consistent results. However, in this operating mode, the rf field is highly non-uniform and frequency shifts in the line centers can result [6, 7]. (See also Fystrom, and Petley and Morris, these Proceedings.) As noted in section 4, in order to investigate the effect of grounding the guard rings for rf (and also the effect of filament current), Sommer et al. carried out a separate series of runs, the results of which are summarized in table 1. A decrease in μ_p/μ_n of about 13 ppm is seen to occur when the guard rings are ungrounded.

We are faced with the question of the significance of this guard ring effect. Unfortunately, it is impossible without a detailed theoretical and experimental investigation to estimate the influence of a non-uniform rf field on the cyclotron resonance line centers. As noted above, the work of Fystrom shows that a significant shift can occur in a non-uniform rf field. Furthermore, Fystrom has also shown that the shape of the cyclotron resonance lines is intimately related to the rf power level and that only for sufficiently large rf powers will the observed line shape approximate the theoretical rectangular line shape. Even when the rf field is uniform, there are shifts in the centers of the resonance lines on going from a high rf power (rectangular) line shape to a low rf power (rounded) line shape. Thus, since Sommer et al. worked in the regime of low rf power as well as non-uniform rf field, the most reasonable approach to this entire problem is to regard the 13 ppm shift as an indication of the possible systematic errors present in the experiment due to these effects rather than as a potential correction. Sufficient information to do otherwise is just not available.

6. Magnetic Field Homogeneity

Another area of interest concerning the Sommer, Thomas, and Hipple experiment is the inhomogeneity of the magnetic field over the active volume of the omegatron. Two relevant magnetic field plots are given in the notebook for the omegatron experiment itself. The first is dated February 23, 1950 and shows a plot of the magnetic field in the vertical or z direction before and after some shimming had taken place. (We use the following coordinate system: x axis = a line normal to and centered on the pole faces; y axis = a line parallel to the pole faces and centered in the magnet gap.) The apparent reduction in field variation was as follows: before shimming, the magnetic field varied by about 90 ppm from $z=-1.0$ cm to $+1.0$ cm ($x=y=0$); after shimming the field varied by only a few ppm over this same distance, although only three points were checked: ± 1.0 cm and the center. It should be noted that the active volume of the omegatron occupied a volume roughly equal to a cube 2 cm on a side centered in the magnet gap.

The second plot is dated July 14, 1950, just prior to the series of 45 final data runs which commenced July 19, 1950 and ended August 15, 1950. It is headed "Magnetic Field Plot" and states "after adjusting the nickel shims in the magnet gap, the following field plot was taken . . . ." An analysis of the 14 numbers comprising the plot shows that the variation in field from $z=-1.0$ cm to $+1.0$ cm ($x=y=0$) was about 9 ppm and from $y=-1.0$ cm to $+1.0$ cm ($x=z=0$) about 6 ppm. However, no data were apparently taken for the x direction.

To give some further indication of the effect of shimming, we point out that Sommer et al. used the same electromagnet as did Thomas, Driscoll, and Hipple in their high-field proton gyromagnetic ratio determination [8]. This magnet was initially

shimmed asymmetrically because sufficiently thin shim stock was unavailable [4]. In this configuration, the field was highly uniform across the bottom edge of the Cotton balance coil, i.e., the y axis. However, the field was highly inhomogeneous along the z axis and along the x axis. Indeed, the magnetic field varied by some 240 ppm from $x=-0.5$ cm to $+0.5$ cm ($y=z=0$) and by some 50 ppm from $z=-1.0$ cm to $+1.0$ cm ($x=y=0$). By contrast the variation from $y=-1.0$ cm to $y=+1.0$ ($x=z=0$) was essentially zero.

During the course of the high field experiment (sometime between January 7 and May 6, 1949, [4, 8]), the magnet was reshimmed symmetrically using thinner shim stock. A field plot dated May 11, 1949 shows the uniformity of the field in the x–y plane ($x=\pm 1/2$ cm, $y=\pm 1$ cm) to be of order 15 ppm, a considerable improvement over the 240 ppm for the asymmetric case.[5] On the other hand, no measurements in the z direction for small x and y (i.e., x and y of order ± 1 cm) are given.

While it is now clear that the field distribution over the active volume of the omegatron as it was used for the 45 final data runs cannot be precisely ascertained, it seems unlikely that Sommer et al. would have ignored a field inhomogeneity as large as several tens of ppm. Indeed, it is difficult to imagine that the influence of such an inhomogeneity on either the proton NMR resonance or the ion cyclotron resonances could have gone unnoticed. Thus, in keeping with the observed magnetic field variations for the different shimming conditions, we are inclined to believe that the field was uniform to something like 10 to 30 ppm over the active volume of the omegatron.

7. Copper Omegatron

For completeness, we summarize the measurements taken by Sommer et al. using a second omegatron. The new omegatron was made of copper (except for a few small pieces of advance, such as the ion collector), and was designed to be electrically identical with the first omegatron when the latter was operated with the guard rings grounded for rf. The second omegatron consisted of a rectangular box, the top and bottom of which formed the rf electrodes. The sides were held at the trapping potential and had suitable apertures for the electron beam. All pieces were gold plated. The tungsten filament required about 4.5 A to obtain the desired electron beam current, 3.5 times the 1.3 A required by the filament used in the first omegatron.

The main function of this second omegatron was to check the 2.6 ppm shielding correction of the first omegatron due to the magnetic properties of the parts used in its construction. The operating characteristics of the two were reasonably similar, although the resolution of the second was somewhat less than that of the first.

[5] The size of the NMR probe used to map the field in both experiments was rather larger than would now be required to obtain the necessary sensitivity. Thus, the maximum excursion possible in the x direction was limited to $\pm 1/2$ cm.

Thirteen runs were made with the copper omegatron from September 7 to September 19, 1950. Then, sometime before October 10, 1950, the tungsten filament was changed so that a filament current of one ampere was all that was required to obtain the necessary beam current. This more closely approximated the situation in the first omegatron. Four additional runs were then made over the period October 12 and 13, 1950. The mean values of μ_p/μ_n for both series of runs were the same and agreed within 1 ppm with the mean of the 45 final data runs.

While it is attractive to try to draw conclusions concerning the shielding correction from these results, we hesitate because of the different filament currents used in the two copper omegatron runs; no shifts in μ_p/μ_n were observed in going from 4.5 A to 1 A as might have been expected, and there is no information available concerning what happened when the filament current was reversed. (It is believed that all of the runs with the copper omegatron were obtained with the filament current in the so called normal direction [4].) There is also the possibility that the variation of the rf field over the active volume was different in the two omegatrons because of the differences in their construction (guard rings vs. solid sides). As noted previously, a non-uniform rf field can significantly affect the position of the cyclotron resonance line centers.

8. Conclusion

The final value for μ_p/μ_n reported by Sommer et al. [1] was

$$\mu_p^*/\mu_n=2.792685\pm 0.000060 \text{ (21 ppm)}, \qquad (1)$$

where the asterisk means for the standard oil NMR probe with no shielding correction to be applied. The stated uncertainty is that originally given by Sommer et al. [1] and was meant to correspond to "several times the estimated probable error." (It is this statement which has no doubt led several constants adjusters [2, 9] to assign a 1 standard deviation uncertainty to the Sommer et al. result of about 10 ppm.) Assuming, as discussed in sections 3 and 4, that the following two corrections hold,

Converting from standard oil sample to spherical water sample	+1.9 ppm
Filament current polarity correction	+7.3 ppm
Total	+9.2 ppm

we obtain

$$\mu_p'/\mu_n=2.792711,$$

where the prime means for protons in a spherical sample of water.

What uncertainty statement can be made concerning this number? Unfortunately, this is a very difficult question to answer at this point in time and with the limited information available. As a first approximation, assume the following 70 percent confidence level "guesstimates" to hold: (1) 13 ppm for possible shifts in the cyclotron resonance line

centers due to a non-uniform rf field and related effects (section 5); (2) 4.1 ppm due to the uncertainty in the filament current correction (section 4); (3) 4 ppm combined uncertainty for shielding of the magnetic field by the omegatron and the measured differences between the standard oil, adjusted water, and pure water NMR probes (2 ppm each, root sum square); and (4) 4 ppm total uncertainty for the statistical error of the mean and the uncertainty in the zero-mass extrapolation used to eliminate line shifts due to electrostatic fields [2]. Add them root sum square to obtain an uncertainty of ± 15 ppm (± 0.000041). However, this does not include the uncertainty due to the inhomogeneity of the magnetic field. As noted, it is difficult to estimate what this might be. If we assume it is equal to the usual one third or so of the total magnetic field inhomogeneity [6, 10], we would assign something like ± 3 ppm to ± 10 ppm. Combining this root-sum-square with the above ± 15 ppm would then yield ± 15 ppm to ± 18 ppm.

In any event, it should now be clear that on the basis of all the information presently available, the ± 21 ppm uncertainty originally listed by Sommer et al. is more nearly equal to a 70 percent confidence interval estimate than has been assumed in the past [2, 9]. Indeed, we would recommend that this ± 21 ppm value be used as the standard deviation uncertainty rather than a lower one in future comparisons of μ_p/μ_n measurements and adjustments of the fundamental constants. We note that by way of this increased uncertainty, the Sommer et al. result (either original or revised) is now in reasonable agreement with the higher values of μ_p/μ_n obtained by Boyne and Franken and Mamyrin and Frantsuzov. Additionally, since the two most recent measurements of μ_p/μ_n, those of Fystrom and of Petley and Morris ([6], these Proceedings), have uncertainties under 10 ppm and agree reasonably well with each other and with the higher μ_p/μ_n values, it seems likely that the higher values are actually the more nearly correct ones, and that a definite conflict with the indirect value determined from the faraday measurement [2] is present.

9. Acknowledgments

We should like to thank P. L. Bender for making the notebooks available to us, and P. L. Bender and D. N. Langenberg for helpful discussions. We very gratefully acknowledge the kind and patient assistance of J. A. Hipple and H. Sommer in trying to unravel these questions some twenty years after the fact. Only classic determinations of fundamental constants remain relevant for such a period of time.

10. References

[1] Sommer, H., Thomas, H. A., and Hipple, J. A., Phys. Rev. **82**, 697 (1951).
[2] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[3] Sommer, H., Ph.D. thesis, 1950, Agricultural and Mechanical College of Texas, unpublished.
[4] Hipple, J. A., and Sommer, H., private communication.
[5] Thomas, H. A., Phys. Rev. **80**, 901 (1950).
[6] Fystrom, D. O., Ph.D. Thesis, 1969, University of Colorado, unpublished; Phys. Rev. Letters **25**, 1469 (1970).
[7] Petley, B. W., and Morris, K., J. Sci. Instr. (J. Phys. E) Series 2, **1**, 417 (1968).
[8] Thomas, H. A., Driscoll, R. L., and Hipple, J. A., J. Res. Nat. Bur. Stand. (U.S.), **44**, 569 (1950).
[9] Cohen, E. R., and DuMond, J. W. M., Rev. Mod. Phys. **37**, 537 (1965).
[10] Petley, B. W., and Morris, K., An Omegatron Measurement of the Magnetic Moment of the Proton in Terms of the Nuclear Magneton, Report QU7, National Physical Laboratory, Division of Quantum Metrology, November, 1969.

DISCUSSION

See the discussion at the end of the paper by Luxon and Rich.

SPECIAL GENERAL INTEREST TALK

Einstein said that tensors could yield
Description of gravity field.
Our scalar Dicke
With experiments tricky
May show that the field is not sealed.

ANON.

Experimental Tests of General Relativity

R. H. Dicke

Joseph Henry Laboratories, Department of Physics, Princeton University, Princeton, N. J. 08540

Philosophical considerations such as those organized under "Mach's Principle" played a central role when Einstein developed general relativity, his tensor field theory of gravitation. But he was mindful of the available observations and made full use of them. First and foremost in helping provide an improved observational foundation for his theory is the modern version of the Eötvös experiment which shows that the gravitational acceleration of a small body is independent of the nature and composition of the body to a part in 10^{11}. More precise measurements utilizing space techniques are possible. The roles of other null experiments will be mentioned. Recent improvements in the observation and interpretation of relativistic effects on planetary orbits and light rays show great promise. The observed solar oblateness may provide a key to the interpretation of Mercury's perihelion rotation. These observations of relativistic orbital effects should eventually be central to the decision between two general relativistic theories of gravitation, the tensor theory and the scalar-tensor theory.

Key words: Eötvös experiment; general relativity; Mach's Principle; scalar-tensor theory; solar oblateness; tensor theory.

DISCUSSION

G. R. Henry: Fairbank and company at Stanford are interested in putting a satellite into orbit around the earth, and, of course, one of the things they hope that they will be able to observe is the Lense-Thirring effect of the earth's rotation. Presumably if you put the same sort of satellite about the sun you might have a shot at seeing the Lense-Thirring effect from the inner core, and you would learn perhaps whether the inner core is in fact rotating or whether there is some strong magnetic field.

R. H. Dicke: I should have mentioned that experiment. It's a very exciting one. Just lack of time to mention all these things. The Lense-Thirring effect is a very interesting effect. It has something to do with Mach's Principle in a way. If you think that inertial masses of objects and inertial coordinate frames are set by the total mass distribution of the universe, then you might well think that a spinning body close by should have some local inertial effects leading to rotation of vertical coordinate frames. In fact, this is a predicted effect of general relativity, an effect expected in the neighborhood of a rotating body. We, in fact, have such a top in space already in the form of the earth-moon system going around the sun. The Lense-Thirring effect is extremely small here because we are so far from the sun. But the rotating sun produces such an effect.

There is another effect that can be measured in addition to Lense-Thirring, the so-called geodesic precession. If you transport a vector parallel to itself—you don't expect it to come back in the same orientation. And there is a tiny effect of that kind on the earth-moon system causing the orbit of the moon to wobble in space as the earth-moon system continues orbiting the sun.

J. Levine: Do you feel that Weber's disk cylinder experiments may shed some light on the scalar-tensor theory?

R. H. Dicke: I hope so. I don't know what you conclude from signals not appearing. There can be many reasons for an apparatus not detecting a signal.

JOSEPHSON EFFECT

Macroscopic Quantum Phase Coherence in Superfluids—Theory*

D. J. Scalapino

Department of Physics, University of California, Santa Barbara, Calif. 93106

Here we review the theory of macroscopic quantum phase coherence and discuss its relationship to measurements of fundamental constants. The off-diagonal matrix elements which characterize this state of matter are introduced and the functional dependence of their phase on the physical constants, potentials, lengths, and times are established. The effects of fluctuations and nonequilibrium conditions on measurements are discussed.

Key words: Fluctuations; fundamental constants; Josephson effects; macroscopic quantum phase coherence.

1. Introduction

In order to appreciate the possibility of using superfluid systems to make high-precision measurements of fundamental constants, it is essential to understand the unique long-range order which characterizes this state of matter. As first recognized by London [1], superfluid He and superconducting metals exhibit quantum mechanical properties on a macroscopic scale. In essence, a macroscopic number of particles occupy the same quantum state. The phase of this quantum state depends upon certain fundamental constants, potentials, distances, and times. Because of the macroscopic occupancy and

* Research supported by the U.S. Army Research Office, Durham, N.C. 27702.

the quantum nature of the phenomena, it is possible to make very accurate observations of relative changes in phase. This determination, along with measurements of the potentials, lengths, and/or times, leads to ratios of the fundamental constants. Here we discuss the present theoretical foundation of these phenomena, emphasizing their relationship to the measurements of the fundamental constants. Our arguments are adapted with these precision measurements [2, 3] in mind; for alternative viewpoints as well as a more comprehensive view, a number of excellent review articles are strongly recommended [4]. In discussing these measurements, it is important to establish (1) the form in which the fundamental constants enter the observations and (2) the effect of fluctuations on these observations.

2. Off-Diagonal Long-Range Order

London [1] was the first to suggest that the superfluid state of matter reflected an underlying quantum behavior. Here we briefly review some simple properties of non-interacting Bose systems and develop the current microscopic formulation of London's suggestion as discussed by Penrose and Onsager [5] and Yang [6].

The ground state of a noninteracting system of bose particles at $T=0$ is

$$|N\rangle = |\varphi_\lambda(1)\varphi_\lambda(2)\cdots\varphi_\lambda(N)\rangle \tag{1}$$

where, for example

$$\varphi_\lambda(x) = (2/V)^{1/2}\sin(\pi x/L). \tag{2}$$

The one-particle density matrix associated with this state can be written in the following forms:

$$\begin{aligned}&\langle N|\Psi^\dagger(x')\Psi(x)|N\rangle\\ &\quad=(2N/V)^{1/2}\sin(\pi x'/L)\,(2N/V)^{1/2}\sin(\pi x/L)\\ &\quad=\langle N|\Psi^\dagger(x')|N-1\rangle\langle N-1|\Psi(x)|N\rangle.\end{aligned} \tag{3}$$

Here $\Psi^\dagger(x')$ is the field creation operator for a particle at point x' and $\Psi(x)$ destroys a particle at x. As the last form indicates, the macroscopic quantum mechanical feature of the state $|N\rangle$ can be expressed as

$$\langle N-1|\Psi(x)|N\rangle = (2N/V)^{1/2}\sin(\pi x/L). \tag{4}$$

This is probably the simplest example of what Yang [6] has called off-diagonal long-range order (ODLRO). The matrix element (4) is off-diagonal in the particle number and it is finite in the limit of a large system where N and $V\to\infty$ but the particle density N/V remains constant. The current density can be written in the form:

$$\begin{aligned}\mathbf{j}(x) = (1/2m)\ \mathrm{Re}&\langle N|\Psi^\dagger(x)|N-1\rangle\\ &\times(\hbar/i)\nabla\langle N-1|\Psi(x)|N\rangle).\end{aligned} \tag{5}$$

This vanishes for the state given by (2) but if, for example

$$\varphi_\lambda = \exp(i\mathbf{p}\cdot\mathbf{x})/V^{1/2}, \tag{6}$$

there would obviously be a current density

$$\mathbf{j} = (N/V)(\mathbf{p}/m). \tag{7}$$

Note that the current depends upon the phase of φ_λ.

Next suppose the temperature is greater than zero but less than the Bose condensation temperature

$$T_c = (2\pi\hbar^2/k_B M)(N/2.6V)^{2/3}.$$

In this case it is convenient to introduce a typical thermal state [7] $|N,\alpha\rangle$. This is a state of N particles and α characterizes the excitations. Specifically, although the boson occupation number

$$\langle N,\alpha|b_p^\dagger b_p|N,\alpha\rangle$$

is definite, this typical thermal state is defined so that

$$\begin{aligned}\sum_p \langle N,\alpha|b_p^\dagger b_p|N,\alpha\rangle F(p)&\\ =\sum_p \frac{1}{\exp(\beta p^2/2m)-1}F(p)&\end{aligned} \tag{8}$$

for all physically interesting $F(p)$. This simple approach can be used within a general elementary excitation picture and simplifies both the notation and the physics to be discussed. For strongly coupled systems a density matrix approach may become more suitable but will not alter any of our conclusions. For $T<T_c$, the non-interacting bose system has ODLRO with

$$\begin{aligned}\langle N-1\alpha|\Psi(x)|N\alpha\rangle&\\ =[1-(T/T_c)^{3/2}]^{1/2}N^{1/2}\varphi_\lambda(x).&\end{aligned} \tag{9}$$

At temperature below T_c, this matrix element is nonvanishing, reflecting the fact that a finite fraction of all the particles is in the lowest single-particle quantum state.

For interacting many-body systems, Penrose and Onsager [5] generalized this simple picture of macroscopic quantum order and suggested that the superfluid state of He can be characterized by the following asymptotic form for the one-particle density matrix.

$$\lim_{|x-x'|\to\infty}\langle N,\alpha|\Psi^\dagger(x')\Psi(x)|N,\alpha\rangle \sim f^*(x')f(x). \tag{10}$$

Here $|N,\alpha\rangle$ represents a typical thermal state of the interacting He system below the lambda point. The ODLRO introduced by Yang [6] corresponds to

$$f(x) = \langle N-1\alpha|\Psi(x)|N,\alpha\rangle. \tag{11}$$

The function $f(x)$ is complex having both an amplitude and a phase. For He it is estimated that $|f^2|$ is of order 0.1.

In order to appreciate the unique feature of the long-range order given by (10) it should be compared with the one-particle density matrix of normal many body systems. The typical range $|x-x'|$ in which the one-particle correlations decay are of order angstroms. For a non-interacting bose system

at a temperature above T_c this range is set by $\hbar/(MkT)^{1/2}$. Thus the long range order represented in (10) is a qualitatively new property of matter.

London's ideas of the inherently quantum mechanical nature of superconductivity were phenomenologically developed by Ginzburg and Landau [8], who introduced a complex order parameter $\tilde{\Psi}$. Following the microscopic theory of superconductivity given by BCS [9], Gor'kov [10] showed that the Ginzburg Landau $\tilde{\Psi}$ was proportional to

$$(V/N)^{1/2}\langle N-2\alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N\alpha\rangle. \tag{12}$$

The ODLRO characteristic of superconductivity is off diagonal in pairs of electrons. While ODLRO arose in the free Bose system because the energy was lower when a finite fraction of the particles occupied the lowest single-particle state, the situation is more subtle in the super-conducting case. BCS showed that the attractive phonon mediated interaction between electrons would lead to a state in which the energy could be lowered if pairs of electrons had the same center of mass wave function. If one calculates the matrix element (12) within the BCS theory, its amplitude is proportional to

$$(N/V)^{1/2}(\Delta/\mu)\ \ln(\omega_0/\Delta). \tag{13}$$

Here Δ is the BCS energy gap, μ is the Fermi energy, and ω_0 is the Debye frequency. Just as in the Bose case, the square of (10) is related to the number of electron pairs which occupy the same center of mass quantum state. At $T=0$ K this fraction is from (10)

$$(N/V)\,(\Delta/\mu)^2 \ln^2(\omega_0/\Delta) \sim 10^{-6} N/V. \tag{14}$$

This of course remains finite as N and $V\to\infty$ with N/V constant. As T approaches T_c it vanishes since Δ goes to zero at T_c. Once again, the superconducting properties are related to this amplitude. Using the Ginzburg-Landau $\tilde{\Psi}$,

$$\mathbf{j}=2e\left[(i\hbar/2m)\,(\Psi\nabla\Psi^*-\tilde{\Psi}^*\nabla\tilde{\Psi})-(2e/mc)\mid\Psi\mid^2\mathbf{A}\right]. \tag{15}$$

3. Forms of the Phase Equations

In discussing the form in which the fundamental constants enter the observations we will focus on the superconducting case. The superfluid He system can be discussed in an analogous manner. We will assume that ODLRO is present, that is, that the matrix element given in (12) is macroscopic, and discuss its phase properties. It is important to realize that this is where the cat goes in the bag (or the rabbit in the hat!) so to speak [11].

First, let's consider how the amplitude given in eq (12) transforms under the gauge change

$$\mathbf{A}\to\mathbf{A}+\nabla\chi,\ \Phi\to\Phi-c^{-1}(\partial\chi/\partial t). \tag{16}$$

Since

$$\Psi(x)\to\Psi(x)\ \exp(ie\chi(x)/\hbar c), \tag{17}$$

it follows that

$$\langle N-2\alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N\alpha\rangle_{A+\Delta\chi}\to\exp(i2e\chi(x)/\hbar c) \times\langle N-2\alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N,\alpha\rangle_A. \tag{18}$$

Note that the factor 2 follows directly from the two particle nature of the superconducting ODLRO. The factors e, h, and c enter as specified by the gauge transformation.

Although this section is concerned primarily with exhibiting the transformation properties of the ODLRO amplitude, it seems worthwhile to look for a moment at a special physical consequence of the form of the transformation (18). Figure 1 shows a multiconnected piece of superconductor with a magnetic field passing through a hole. For fields less than the critical field, B vanishes deep inside the metal so that its effect can be represented by χ. Here the depth scale is set by the field penetration depth of order hundreds of angstroms. In this case

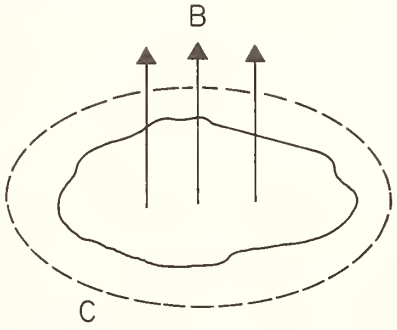


FIGURE 1. *Multiconnected piece of superconductor with a magnetic field passing through a hole.*

χ is multivalued since

$$\oint\nabla\chi\cdot dl=\oint B\cdot dS. \tag{19}$$

Now, if the ODLRO amplitude is to remain single-valued then the exponential factor in (18) must have the same value regardless of how the point x is approached. In going around the hole along a path C which is deep inside the metal, this single valuedness will only be possible if

$$\frac{2e}{\hbar c}\oint_C A\cdot dl=2\pi n. \tag{20}$$

This implies that the flux contained in C is quantized in units of $hc/2e$. The observations of flux quantization [12] in 1961 by Denver and Fairbank and Doll and Nabauer were the first experiments to focus directly on the unique phase coherent properties of the superconducting state. In principle measurements of this type could provide information on hc/e. At present, difficulties in measuring areas and magnetic fields have restricted these measurements to parts in 10^3 [13]. To this accuracy, the flux quantum is $hc/2e$.

Next, consider in thermal equilibrium the time development of the amplitude

$$\langle N-2\alpha \mid \exp[i(Ht/\hbar)]\Psi_\uparrow(x) \times \Psi_\downarrow(x)\exp[-i(Ht/\hbar)] \mid N,\alpha\rangle = \exp[-i(E_{N,\alpha}-E_{N-2,\alpha})t/\hbar] \times \langle N-2\alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N\alpha\rangle. \quad (21)$$

Here $E_{N,\alpha}$ is the energy of a typical thermal state with N electrons while $E_{N-2,\alpha}$ has the same set of excitations but 2 less electrons. Since the α quantum numbers are not changed, the entropy is constant as well as the volume and

$$E_{N,\alpha}-E_{N-2,\alpha}=2(\partial E/\partial N)\mid_{S,V}=2\mu. \quad (22)$$

Here μ is the electrochemical potential. For a system with μ slowly varying compared to the coherence length ξ this becomes $2\mu(x)$. Thus the time development of the amplitude is proportional to $\exp[-i2\mu t/\hbar]$ where μ is the electrochemical potential. This form was first discussed by Gor'kov

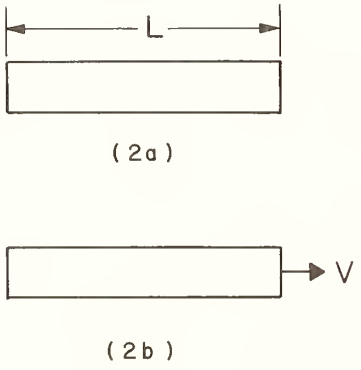


FIGURE 2a. *Superconducting strip of length* L *at rest in laboratory*, 2b. *Superconducting strip moving with velocity* v *relative to laboratory.*

[10] and later given great emphasis by Josephson's [14] work.

As a final example of the form of the equations which determine the phase transformation properties of the ODLRO amplitude, consider the superconducting strip shown in figure 2. Suppose when it is at rest with respect to the lab frame (2a) the difference in the phase of the $\langle N-2\alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N\alpha\rangle$ amplitude between the ends is $\Delta\varphi$. How does this change if it is in motion? Non-relativistically this change in state can be generated by the Galilean transformation

$$U=\exp(iM\mathbf{v}\cdot\mathbf{R}/\hbar). \quad (23)$$

Here M is the total mass of the system, $\mathbf{v}$ is the velocity of the strip and $\mathbf{R}=M^{-1}\sum_i M_i\mathbf{x}_i$ is the center of mass operator for the strip.

$$\langle N-2\alpha \mid U^+\Psi_\uparrow(x)\Psi_\downarrow(x)U \mid N\alpha\rangle=\exp(i2m\mathbf{v}\cdot\mathbf{x}/\hbar) \times \langle N-2\alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N\alpha\rangle. \quad (24)$$

Therefore the change in the phase between the ends for the case shown in (2b) is $\Delta\varphi+(2m/\hbar)vL$. In this non-relativistic situation the phase shift depends upon the free mass of the electron. Since the electrons are in fact bound by the work function $\tilde{\varphi}$, Josephson [15] has suggested that m should in fact be replaced by

$$m^*=m[1-(\tilde{\varphi}/mc^2)]. \quad (25)$$

For typical metals this correction is of order 10^{-5}.

4. The Josephson Effect

In section 3 we saw that the phase of the ODLRO matrix element under gauge, time, and Galilean transformation involved potentials such as $\mathbf{A}$ and μ, lengths and times and the fundamental constants e/hc, $1/h$ and m^*/h. Although flux quantization [12] was the first really macroscopic quantum effect of this class which was observed, it was Josephson's theory of pair tunneling and the subsequent observation of the effect by Anderson and Rowell [16] which showed how one could study relative phase

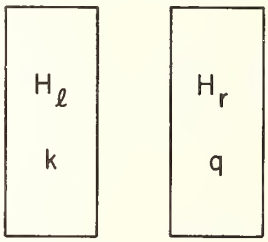


FIGURE 3. *Schematic illustration of a tunnel junction.*

changes in the ODLRO matrix element. Here we will briefly review some of the important features of this effect taking Anderson's point of view which emphasizes the role of the relative phase dependence of the coupling energy of the tunnel junction. This approach was first introduced in [16] and is discussed by Anderson in the Ravello Lectures [4] and Ferrell and Prange [17].

Josephson considered the problem of the establishment of ODLRO between two superconductors. Figure 3 shows a schematic illustration of a tunnel junction in which the left hand superconductor is described by a Hamiltonian H_l and its electron states are labeled by k while the righthand superconductor H_r has its electron states labeled by q. When the two superconductors are well separated, tunneling is negligible and the Hamiltonian simply consists of H_l+H_r. In this case, neglecting the small Coulomb interaction, the states

$$\mid \Psi_n\rangle = \mid N-2n\rangle_l \mid N+2n\rangle_r \quad (26)$$

in which n pairs have been transferred from the left to the right hand superconductor are degenerate. As the superconductors are brought closer together

(~10 A), the tunneling interaction

$$H=\sum_{kq} T_{kq}C_k^{+}C_q+\text{h.c.} \tag{27}$$

lifts this degeneracy. It has a finite matrix element, in lowest order, of the form

$$\langle\Psi_{k\pm1}\mid H_T(E_0-H_1-H_r)^{-1}H_T\mid\Psi_n\rangle=-\tfrac{1}{2}E_1. \tag{28}$$

Here for identical superconductors having a single-single particle electron density of states, $N(0)$, an average tunneling matrix element $\mid T\mid^2$, and a gap Δ,

$$E_1=\pi^2N^2(0)\mid T\mid^2\Delta \tag{29}$$

which vanishes above T_c. We will see shortly how ODLRO explicitly enters in giving a nonvanishing E_1. The linear combination of $\mid\Psi_n\rangle$ states which diagonalizes this interaction is

$$\mid\Psi_\varphi\rangle=\sum_n e^{in\varphi}\mid\Psi_n\rangle. \tag{30}$$

This state has an energy

$$E=E_0-E_1-E_1\cos\varphi \tag{31}$$

which is a minimum when φ is 0 (mod 2π). Furthermore, a current flows between the two superconductors if $\varphi\neq0$ (mod 2π). It is given by

$$I=(2e/\hbar)(\partial/\partial\varphi)E=I_1\sin\varphi \tag{32}$$

with

$$I_1=(2e/\hbar)E_1. \tag{33}$$

Now, where is the notion of ODLRO and the possibility of using this tunneling phenomena to probe the relative phase changes discussed in section 3? It has been buried in eq (28) and (32). Writing out the matrix elements in detail one finds that

$$I\propto\text{Im}(e^{i\varphi_r}\langle N+2(n+1),\alpha\mid C_q^{+}C_{-q}^{+}\mid N+2n,\alpha\rangle_{rl}$$
$$\times\langle N-2(n+1),\alpha'\mid C_kC_{-k}\mid N-2n,\alpha'\rangle_l). \tag{34}$$

Here sums and energy denominators have been left out in order to ease the notation. They do not affect the phase dependencies which we are interested in exhibiting. It is clear from (34) that the relative phase changes in the amplitude discussed in section 3 directly enter in (34). The actual magnitude of this part of the current depends on many details of both material and geometry but this magnitude is not of interest, other than that it be observable and sufficiently large (as we will discuss in the next section) to prevent fluctuations which could obscure the observation of the phase.

Now consider, for example, the time dependence of I. Inserting the time translation operator simply generates as previously discussed a factor $\exp(i2\mu_rt/\hbar)$ and $\exp(-i2\mu_lt/\hbar)$ for each of the ODLRO elements in (34) giving an over all factor

$$\exp[i2(\mu_r-\mu_l)t/\hbar]=\exp[i2\Delta\mu t/\hbar] \tag{35}$$

inside the brackets in eq (34). Therefore

$$I=I_1\sin[(2\Delta\mu t/\hbar)+\varphi] \tag{36}$$

and the current oscillates at a frequency set by the difference in electrochemical potential between the two superconductors. If, in addition, a microwave field of frequency ω is coupled to the junction it introduces an oscillating phase change via χ [eq (18)] so that

$$I=I_1\sin[(2\Delta\mu t/\hbar)+\alpha\sin\omega t+\varphi]. \tag{37}$$

This will have a dc part whenever

$$2\Delta\mu/\hbar=n\omega. \tag{38}$$

By synchronizing the rate of the relative ODLRO phase change across a junction at a multiple $n\omega$ of the applied microwave field, an electrochemical potential $\Delta\mu$, given by the Josephson condition (38), is established. Then, the property of infinite conductivity of the superconducting links faithfully transmits this electrochemical potential difference to normal leads where a null potentiometer measurement compares it with the electrochemical potential of a standard cell. This latter potential can be calibrated in absolute volts times the free electron charge. It is important to keep clearly in mind that it is only in this final calibration step that e enters. The superfluid aspect of the measurement relates ω to $2\Delta\mu/\hbar$. Thus, questions about possible renormalization of e in metals are actually irrelevant [18].

5. Fluctuations

In section 3 the equilibrium form of the phase equations was discussed. In an actual measurement, the effect of fluctuations and deviations from equilibrium must be considered. Roughly the fluctuations can be divided into quasi-particle and order-parameter fluctuations. The former provide a drive for the latter which directly affect the phase measurements. However, the order-parameter fluctuations involve an activation energy which can be made sufficiently large, in a properly designed experiment, that their effect can be made negligible. The effects of deviations from thermal equilibrium are not as well understood but, generalizing from some specific cases, two important features emerge: the quantum nature of the long-range order occurs for a range of non-equilibrium states and the phase information can be faithfully propagated from a non-equilibrium to an equilibrium part of the super-system. Here we will briefly review the mechanisms that give rise to these effects in the e/h measurement and summarize with references to the literature what is presently known.

The tunneling interaction H_T can not only transfer pairs but at finite temperatures it gives rise to quasi-particle tunneling. In lowest order this adds additional quasi-particle and quasi-particle-pair currents [4]

$$I_{qp}(V)+\sigma_1V\cos\varphi \tag{39}$$

to the pair current eq (32). In higher order there are further modifications of the quasi-particle current

$I_{qp}(V)$ and additional $\cos n\varphi$ and $\sin n\varphi$ contributions as well. Generally these higher order terms are weak for tunnel junctions. In addition there are quasi-particle current noise fluctuations [19, 20]. Finally, when the self-radiation field of the junction is included, besides giving rise to the well known Fiske steps [21], there are in addition pair fluctuation noise [22, 20] and frequency pulling effects [22, 23].

For the purposes of this discussion it is useful to separate the noise fluctuations from the nonthermal equilibrium quasi-particle current and frequency pulling effects. The effect of the noise fluctuations were qualitatively discussed by Anderson [4] and Scalapino [19] who pointed out the importance of the activation energy barrier to phase slippage. The energy for a tunnel junction eq (31) changes by $\hbar I_1/2e$ as the phase φ varies by 2π. Thus for a thermal noise fluctuation to introduce a fluctuation 2π in phase it must overcome this activation energy. Recently, detailed calculations [24, 25] of the effect of noise fluctuations on the step width in the e/h measurement have confirmed this simple picture. The voltage width ΔV of a step occurring at V is approximately given by

$$\Delta V/V \sim \exp[-\hbar I_s/2ekT] \qquad (40)$$

where I_s is the radiation induced step height. For typical operating conditions this can be made negligible compared to the experimental resolution.

The problem of nonequilibrium effects is not yet so clearly answered. However, a very important contribution to an understanding of this problem was recently made by McCumber [26]. Stephen [22] and Scully and Lee [23] had predicted that the radiation emitted from a Josephson junction would be slightly (~ 1 in 10^8) shifted from the Josephson frequency $2eV/h$. This frequency pulling effect arose from coupling the junction to a radiation field. However, as McCumber explicitly calculated, the electrochemical potential difference is also modified by the presence of the radiation field. This modification is just such that the Josephson frequency relation is exactly satisfied. The point again is that the frequency is set by $2\Delta\mu/h$ and eV simply serves to convert $\Delta\mu$ to voltage. Here, the driven radiation field corresponds to a nonthermal equilibrium situation, μ is shifted, but the basic phase relation persists.

Josephson [4] pointed out that a local departure from equilibrium could modify the time phase relation so that

$$(\partial/\partial t)\ \arg\langle N-2, \alpha \mid \Psi_\uparrow(x)\Psi_\downarrow(x) \mid N\alpha\rangle = -2(\mu(x,t)+\mu'(x,t))/\hbar. \qquad (41)$$

Here μ' depends upon the departure from equilibrium. McCumber's calculation of the shift μ' due to the radiation field is an example of this phenomena. The same type of situation may arise due to local quasi-particle transfer. This will not however affect the measurement if (1) the superconducting leads remain resistanceless (i.e., no flux motion) and (2) the contacts of the normal leads at which the null potential reference is made are at points of thermal equilibrium. Work remains to be done on the quantitative aspects of this last condition but the characteristic length is probably a thermal mean free path.

6. Acknowledgment

The author wishes to acknowledge helpful discussions with J. Clarke, J. B. Hartle, D. N. Langenberg, W. H. Parker, J. R. Schrieffer, and B. N. Taylor, He would also like to express his appreciation for the hospitality of the University of Pennsylvania where this manuscript was assembled.

7. References

[1] London, F., Proc. Roy. Soc. **A152**, 24 (1935).
[2] Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177**, 639 (1969), and this Conference.
[3] Parker, W. H., and Simmonds, M., this Conference; Khorana, B. M., and Douglass, D. H., Jr., this Conference.
[4] Anderson, P. W., 1963 Ravello Summer School Lectures on the Many-Body Problem, ed. by E. R. Caianiello (Academic Press, Inc., New York, 1964), Vol. II; Josephson, B. D., Adv. in Phys. **14**, 419 (1965); Bardeen, J., article in Quantum Theory of Atoms, Molecules, Solid State, (Academic Press, Inc., New York, 1966); Anderson, P. W., article in Progress in Low Temperature Physics, ed. C. J. Gorter, (North-Holland Publ. Co., Amsterdam, 1967); Clarke, J., to be published, Am. J. Phys.
[5] Penrose, O., and Onsager, L., Phys. Rev. **104**, 576 (1956).
[6] Yang, C. N., Rev. Mod. Phys. **34**, 694 (1962).
[7] J. Bardeen seems to have introduced this useful notion; see Schrieffer, J. R., Theory of Superconductivity, (W. A. Benjamin, New York, 1964), p. 62.
[8] Ginzburg, V. L., and Landau, L. D., Zh. Eksperim. i. Teor. Fiz. **20**, 1964 (1950).
[9] Bardeen, J., Cooper, L. N., and Schrieffer, J. R., Phys. Rev. **108**, 1175 (1957).
[10] Gor'kov, L. P., JETP **1**, 505 (1958).
[11] There are in fact numerous derivations of the types of relationships we will discuss. They differ primarily in the manner in which the cat is put in the bag. For example Byers, N., and Yang, C. N. (Phys. Rev. Letters **7**, 46 (1961)) used gauge invariance and single valuedness of the many-body wave function to show that if the fluxoid were quantized, the units would be hc/ne with n an integer. The "if" is dependent of course on the existence of the long range superconducting order which in addition selects n equal to 2. This latter feature is, from an experimental point of view, irrelevant since needless to say present measurements are more precise than a factor of 2. Recently, Bloch, F. (Phys. Rev. Letters **21**, 1241 (1968)) has extended these arguments and within the additional limitation of reversibility has discussed the Josephson frequency condition from this point of view.
[12] Deaver, B. S., Jr., and Fairbank, W. M., Phys. Rev. Letters **7**, 43 (1961); Doll, R., and Nabauer, M., Phys. Rev. Letters **7**, 51 (1961).
[13] Goodman, William L., and Deaver, Bascom S., Jr., Phys. Rev. Letters **24**, 870 (1970).
[14] Josephson, Phys. Letters **1**, 251 (1962), and references listed in [4].
[15] See comment by Anderson, P. W., reference [4].
[16] Anderson, P. W., and Rowell, J. M., Phys. Rev. Letters **10**, 230 (1963).

[17] Ferrell, R. A., and Prange, R. E., Phys. Rev. Letters **10**, 479 (1963).
[18] Nordtvedt, K., Jr., Phys. Rev. **B1**, 81 (1970) has argued that quantum electrodynamic corrections will result in shifts of the e/h ratio observed via the Josephson effect measurement of reference [2]. This is not the case as pointed out in comments to be published in Phys. Rev. by Langenberg, D. N., and Schrieffer, J. R., and Hartle, J. B., Scalapino, D. J., and Sugar, R. L.
[19] Scalapino, D. J., Proc. Symp. on the Physics of Superconducting Devices, Univ. of Virginia, April 1967 (ed. by B. S. Deaver and W. S. Goree).
[20] Dahm, A. J., Denenstein, A., Langenberg, D. N., Parker, W. H., and Scalapino, D. J., Phys. Rev. Letters **22**, 1416 (1969).
[21] Fiske, M. D., Rev. Mod. Phys. **36**, 221 (1964).
[22] Stephen, M. J., Phys. Rev. Letters **21**, 1629 (1968).
[23] Scully, M. O., and Lee, P. A., Phys. Rev. Letters **22**, 23 (1969).
[24] Stephen, M. J., to be published, Phys. Rev.; Waldram, J. R., Pippard, A. B., and Clarke, J., to be published, Proc. Roy. Soc.
[25] Kose, V. E., and Sullivan, D. B., J. Appl. Phys. **41**, 169 (1970).
[26] McCumber, D. E., Phys. Rev. Letters **23**, 1228 (1969).

DISCUSSION

D. J. Scalapino: I meant to add that in all the time that I've been involved in any of this I have been amused by the fact that people have always asked me, "Is it 2 or is it e?" But no one ever asks, "Is it h?" And I don't know why that is. It's maybe $\hbar = 1$ or maybe we're uncertain of h and go backwards to 2π. But the 2 and the e have taken a lot of beating, and maybe it's time we questioned the h. (*Laughter.*)

Said V. Hughes at a conference luncheon,
"The fine structure constant's a function
Of levels in muonium
And perhaps positronium,
But not any Josephson junction!"

S. L. Kaufman

Said D. Langenberg to V. Hughes,
"To each his particular views.
But so long as my dewar
Makes ppms fewer
I'll follow Josephson's clues."

D. N. Langenberg

The Josephson Effect as a Basis for Determination of Fundamental Constants—Experiment*

D. N. Langenberg

Department of Physics, University of Pennsylvania, Philadelphia, Pa. 19104

The Josephson effect in systems of weakly coupled superconductors or superfluids has recently assumed considerable importance in the determination of several fundamental physical constants. It already provides the basis for the most accurate determination of e/h and promises to yield values of h/m_e and h/M_{He} of significant accuracy in the near future. The experimental properties of the Josephson effect which are important in such determinations are discussed. Achievements to date are reviewed, and possible future developments are described.

Key words: Fundamental physical constants; Josephson effect; superfluids; weakly coupled superconductors.

1. Introduction

In 1962, B. D. Josephson discovered theoretically some remarkable consequences of macroscopic quantum phase coherence in superconductors in circumstances where two superconductors are weakly coupled together [1]. Analogous effects have since been observed in weakly coupled superfluid helium systems. These Josephson effects provide a basis for the determination of several fundamental physical constants with accuracies comparable to or greater than those obtainable using more conventional methods. In the previous paper Scalapino discusses the state of our theoretical understanding of macroscopic quantum phase coherence and the Josephson effects. An appreciation of the strengths and limitations of these basic concepts is essential in any critical assessment of a fundamental constant determination based on the Josephson effects. Just as essential is an appreciation of the strengths and limitations of the experimental technology which is required for the exploitation of the Josephson effects in precision measurements. In the present paper we describe some of the experimental manifestations of the Josephson effects, show how they can be used to determine certain fundamental constants, review the progress which has already been made in such determinations and in precision testing of the validity of the theory, and outline some possibilities for future developments. Our objective is to set the scene for the following papers, in which current workers in the field give the reader a view of the state of the art in mid-1970. We begin by discussing superconductor Josephson junctions, which play a key role in the fundamental constant determinations so far reported.

2. The Josephson Junction

A Josephson junction can be viewed rather generally as a "weak" spot in a superconducting system at which a localized quantum phase difference can be established between two superconducting regions and monitored by observing the phase-dependent supercurrent through the junction. The response of this junction supercurrent to electric and magnetic fields leads to a variety of interesting and useful Josephson phenomena. The most important types of junction in which the necessary weak coupling has been achieved are shown in figure 1. The type originally considered by Josephson is the tunnel junction

* Supported by the National Science Foundation and the Advanced Research Projects Agency.

shown in figure 1(a). Here two superconducting films are coupled via phase coherent tunneling of bound Cooper pairs through an insulating barrier between the films. The films are commonly formed by vacuum evaporation. The barrier is usually an oxide formed by exposing the first evaporated film to an oxygen atmosphere or to a glow discharge in oxygen. Figure 2(b) shows a Dayem bridge [2] or weak link in which two parts of a single superconducting film are coupled through a small constriction. If such a device is to be sufficiently weakly coupled to exhibit Josephson behavior, the dimensions of the constriction must be comparable with or smaller than the coherence length of the superconductor. This length is typically a few thousand angstroms, so that the

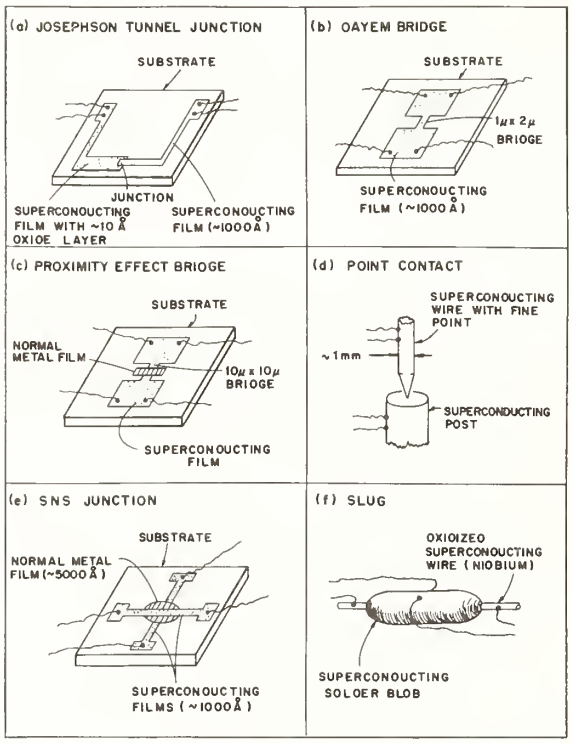


FIGURE 1. *The most common types of Josephson junctions.*

required dimensions are near the limits achievable by mechanical scribing, evaporation through or around masks, or photoresist techniques. Various successful applications of these techniques have been developed. The coherence length is temperature dependent and increases rapidly as the superconducting transition temperature T_c is approached from below. Consequently, Dayem bridges are often usable only in a restricted temperature range just below T_c. The severity of the restriction on link dimensions is considerably lessened in the proximity effect bridge [3] shown in figure 1(c). Here there is a normal metal over- or underlay at the bridge. The normal metal "weakens" the superconductivity of the bridge through the proximity effect, and the effective coherence length of the superconductor is increased just as it is when the superconductivity is weakened near T_c. The bridge can therefore be larger and easier to fabricate. The point contact shown in figure 1(d) is extremely simple to fabricate and its characteristics can be varied by adjusting the contact pressure [4]. Its properties usually are similar to those of the several types of bridges, although with care tunneling (presumably through an oxide barrier at the contact) can be observed. The superconductor–normal metal–superconductor (SNS) junction shown in figure 1(e) consists of two superconducting films separated by a relatively thick (1000 to 10 000 Å) layer of normal metal [5]. A weak superconductivity is induced in the normal metal by the superconductors (the proximity effect again) and this leads to the necessary phase coherent coupling of the superconductors. The "slug" shown in figure 1(f) is formed simply by freezing a blob of ordinary solder around an oxidized niobium wire [6]. Multiple Josephson junctions are formed at the niobium-solder interface, probably by small bridges through the oxide. The resulting device displays rather complex behavior, but it has been used successfully in several Josephson junction applications.

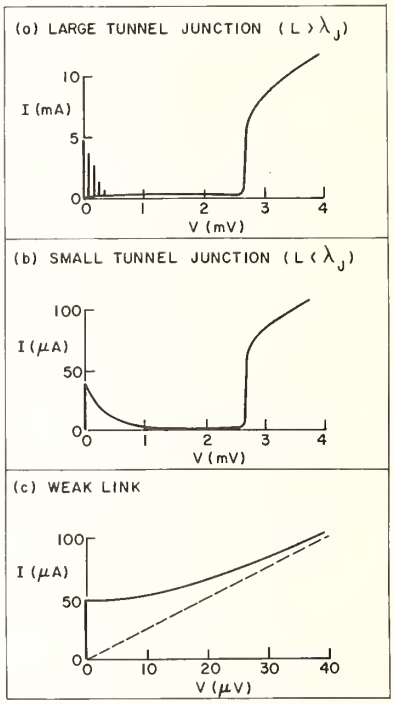


FIGURE 2. *Typical Josephson junction current-voltage characteristics.*
λ_J here is the Josephson penetration depth [7].

The behavior of the supercurrent in these junctions is described by the basic Josephson equations [7]

$$j = j_1 \sin(\varphi + 2e \int \mathbf{A} \cdot d\mathbf{r}), \quad (1)$$

$$\partial\varphi/\partial t = 2\Delta\mu/\hbar, \quad (2)$$

where φ is the quantum phase difference across the junction, j is the supercurrent density, $\mathbf{A}$ is the mag-

netic vector potential, and $\Delta\mu$ is the electrochemical potential difference across the junction. The parameter j_1 depends on the junction material and temperature [8] and also has an energy or voltage dependence [9]. The latter can often be ignored and j_1 treated as a constant for a given junction at fixed temperature. The most important contribution to the electrochemical potential is the electrostatic potential eV, and eq (2) is often written

$$\partial\varphi/\partial t = 2eV/\hbar.$$

For $V = \mathbf{A} = 0$, eq (1) and eq (2) predict a zero-voltage (resistanceless) supercurrent density with maximum value j_1. For $V \neq 0$, they predict an oscillating (ac) supercurrent with characteristic frequency $\nu = 2\Delta\mu/h = 2\ \mathrm{eV}/h$. This is the Josephson frequency-voltage relation, which provides the key to the application of the ac Josephson effect to the determination of $2e/h$. Its validity depends directly on the validity of eq (2). The theoretical arguments for and against belief in the ultimate exactness of eq (2) have been discussed by Scalapino in the previous paper. The experimental evidence is reviewed in the next section. The simple sinusoidal form of eq (1) appears to account rather well for almost all observed Josephson phenomena. However, it probably does not provide an exact description of the current-phase relationship, particularly in the various types of weak links, nor is its exactness essential to fundamental constant determinations based on the Josephson effects.

The essential nonlinearity of eq (1) leads to a rich variety of Josephson phenomena which we have not the space to discuss here [7]. However, we show in figure 2 dc current voltage (I-V) characteristics typical of the several types of Josephson junctions. The sharp increase in current at 2.7 mV in figures 2(a) and 2(b) is due to tunneling of single quasiparticles (Giaever tunneling [10]) and is not relevant to the Josephson effect. The zero-voltage supercurrent ranging up to a maximum (critical) value is apparent in all three curves. The current component due to flow of normal quasiparticles at small nonzero voltages is negligible in tunnel junctions but is important in weak links where it appears as an ohmic background current. There is in addition a dc supercurrent at nonzero voltages; in small tunnel junctions and weak links (figs. 2(b) and 2(c)), it decreases monotonically with increasing voltage and is simply the dc component of the periodic but not sinusoidal ac supercurrent [11]. A tunnel junction forms a short section of parallel plate transmission line and hence can support resonant electromagnetic modes [12]. In large tunnel junctions, resonant coupling of the ac supercurrent to these modes produces the singularities (Fiske steps) shown in the I-V characteristic of figure 2(a) [13].

3. Determination of $2e/h$—The Josephson Voltage Standard

It is obvious that if an electrochemical potential difference $\Delta\mu$ is established across a Josephson junction and the frequency ν of the resulting ac Josephson supercurrent is measured, the ratio $\nu/\Delta\mu$ yields a value of $2/h$. If, as is conventional, the electrochemical potential is specified in terms of the equivalent electrostatic potential V through $\Delta\mu = eV$, the ratio ν/V yields a value of $2e/h$. There are two experimental manifestations of the ac supercurrent which can be exploited in such an experiment. The ac supercurrent in a voltage-biased junction emits coherent electromagnetic radiation [14]. Measurements of the frequency of this radiation and of the bias voltage can and have been used to determine $2e/h$ [15]. However, the low power level of the radiation (typically 10^{-11} W) and the limited size of the usable bias voltage make a second effect more convenient in practice. If a Josephson junction is exposed to an rf field (usually at microwave frequencies), well-defined current steps appear in the dc I-V characteristic at a series of discrete voltages $V_n = nh\nu/2e$, where ν is the source frequency and n is an integer [16]. An example is shown in figure 3. These Shapiro

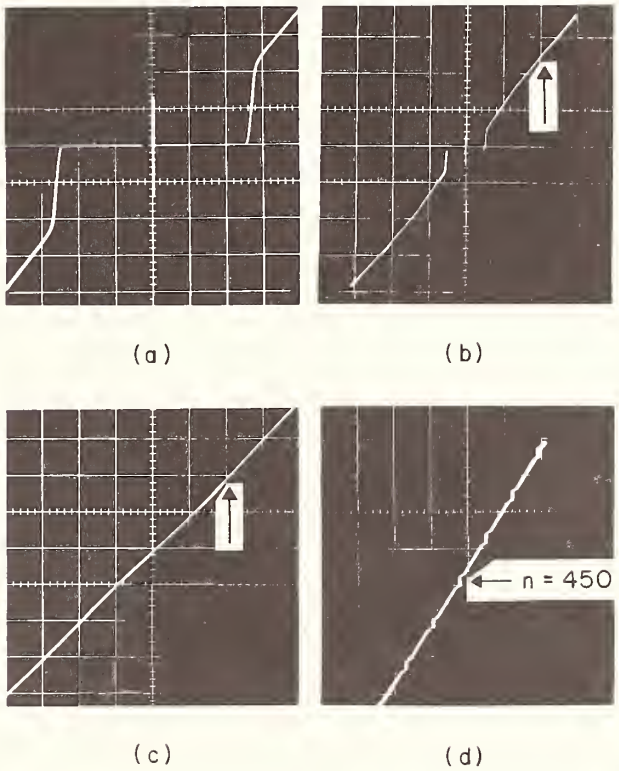


FIGURE 3. (a) *The current-voltage characteristic of a* Pb–PbO_x–Pb *Josephson tunnel junction.*
Current is vertical with a scale of 2.5 mA/cm; voltage is horizontal with a scale of 1 mV/cm.
(b) *The same characteristic with current scale 10 mA/cm and voltage scale 5 mV/cm.*
The arrow just above 10 mV marks the site of (d).
(c) *The characteristic with 11 GHz microwave radiation incident.*
The scales and the arrow are the same as in (b).
(d) *Microwave-induced Shapiro steps above 10 mV.*
The current scale is 50 μA/cm and the voltage scale is 25 μV/cm.

steps result from the beating of the ac supercurrent with harmonics of the rf current induced in the junction by the incident radiation. This effect has been used in all of the accurate ac Josephson effect determinations of $2e/h$ to date.

A determination of $2e/h$ by this method reduces to a measurement of a frequency and a corresponding dc voltage. Microwave frequencies can be measured routinely with accuracies of the order of 0.01 parts per million (ppm) or better using off-the-shelf commercial equipment. The voltage measurement is much more difficult and less accurate, and therefore determines the overall accuracy of the experiment. The first of a series of determinations of $2e/h$ reported by the University of Pennsylvania group was carried out using an available and rather aged commercial potentiometer and had a quoted uncertainty of 60 ppm, limited by the potentiometer calibration [17]. The second determination [15, 18] used a state-of-the-art commercial potentiometer and had a quoted one-standard-deviation uncertainty of 2.4 ppm, again attributable primarily to potentiometer calibration uncertainties. This determination had a number of significant consequences. Together with other recent experiments, it resolved an important discrepancy between existing experimental values of the fine structure constant. This led to a reevaluation of the fundamental physical constants and a critical review of the state of agreement between theory and experiment in quantum electrodynamics by Taylor, Parker, and Langenberg [19]. This determination also demonstrated that a Josephson junction has great potential as a maintenance standard or even a primary or defining standard of emf [20]. This latter possibility has stimulated investigations of the method in a number of laboratories. Petley and Morris have reported a determination of $2e/h$ at the National Physical Laboratory (United Kingdom) with a quoted uncertainty of 2.2 ppm [21]. Their result is in excellent agreement with the value reported by Parker, Langenberg, Denenstein, and Taylor (hereafter called PLDT) [15] and also with the revised version of the PLDT result recently reported by Denenstein et al. [22]. Taken together, these measurements provide a value of $2e/h$ with a one-standard-deviation uncertainty of about 1.5 ppm, which fits rather neatly into the present overall framework of our knowledge of the fundamental constants [19].

It has now been less than four years since the first really high-accuracy ac Josephson effect determination of $2e/h$ was reported. It is reasonable to suppose that the method has not yet been developed to its practical limits in this relatively short time. We may therefore ask how far it is feasible to push it, what would be the impact of further development on our knowledge of the fundamental constants and on voltage standard technology, and, last but not least, to what extent can we rely on the fundamental quantitative validity of the Josephson frequency-voltage relation on which the method rests.

First let us consider the relationship between $2e/h$ and the other fundamental constants, and between the "Josephson volt" and the conventional electrochemical volt. In practice, the volt which enters an ac Josephson effect determination of $2e/h$ is not the absolute volt, but a "local volt" provided by a reference electrochemical standard cell located in the laboratory in which the experiment is done. This in turn can be compared with some central standard, e.g., the primary reference group of standard cells maintained by the United States National Bureau of Standards (NBS). The result of a specific $2e/h$ determination must thus be referred to a volt *as maintained by a specified laboratory at a specified epoch*. Now the uncertainty in the relationship between a typical as-maintained volt and the absolute volt is currently about 3 ppm [19]. There is thus an unavoidable additional contribution to the uncertainty of $2e/h$ in absolute units, which is in fact larger than the uncertainties of the determinations mentioned above. This difficulty is common to all fundamental constant determinations involving measurements of current or voltage. It is most appropriately dealt with by including an as-maintained to absolute-ampere ratio among the constants to be adjusted as was done by Taylor, Parker, and Langenberg in their recent adjustment of the constants [19].

There is an important case involving $2e/h$, however, where the experimental uncertainty of $2e/h$ in terms of an as-maintained volt is the uncertainty of primary interest. That most important constant, the fine structure constant, can be related to $2e/h$ by

$$\alpha^{-2}=\left(\frac{1}{4R_\infty}\frac{c\Omega_{\mathrm{ABS}}}{\Omega_{\mathrm{NBS}}}\frac{\mu_p'}{\mu_B}\right)\left(\frac{2e/h}{\gamma_p'}\right), \tag{3}$$

where R_∞ is the Rydberg constant, c is the velocity of light, $\Omega_{\mathrm{ABS}}/\Omega_{\mathrm{NBS}}$ is the ratio of the absolute ohm to the ohm as maintained by NBS, μ_p'/μ_B is the proton magnetic moment in Bohr magnetons, and γ_p' is the proton gyromagnetic ratio. The primes indicate determinations in spherical samples of water, and it is assumed that the two quantities in the second parenthesis are determined in terms of NBS as-maintained electrical units. Because $2e/h$ is experimentally the ratio of a frequency and a voltage, and γ_p' is the ratio of a frequency and a magnetic field, which for a low-field determination is the product of a coil geometric constant and a current, the rather uncertain as-maintained-ampere-to-absolute-ampere ratio cancels out of eq (1), leaving the absolute-to-as-maintained-ohm ratio indicated. The quantity $c\Omega_{\mathrm{ABS}}/\Omega_{\mathrm{NBS}}$ has an experimental uncertainty of about 0.4 ppm, and the uncertainties of R_∞ and μ_p'/μ_B are 0.1 and 0.07 ppm respectively. A typical uncertainty for a low-field determination of γ_p' is 4 ppm. We see, therefore, that the uncertainty of a value of α determined indirectly through eq (1) is controlled by the uncertainties of the two quantities in the second parenthesis, $2e/h$ and γ_p'. The most accurate values of α presently available can in fact be obtained by this indirect route, using the $2e/h$ values of PLDT, Petley and Morris, or Denenstein et al., and a combination of existing γ_p' measurements. The uncertainties in the resulting values of α are about 2 ppm.

A reduction of the uncertainty of a $2e/h$ determination to *zero* would lower the uncertainty in α only to about 1.5 ppm. From this discussion it is apparent that the real importance of greatly improved accuracy in the ac Josephson effect determination of $2e/h$ is the opportunity it affords to obtain a significantly improved accuracy in α by improving the determination of γ_p'. Given a $2e/h$ with an uncertainty of several tenths of a ppm, the uncertainty of α would be completely determined by the uncertainty in low-field determinations of γ_p'. Any improvement in the latter would yield an immediate corresponding improvement in our knowledge of α. There is an exciting experimental challenge here! It is not obviously impossible that a complete rethinking of the low-field γ_p' experiment might lead to an order of magnitude improvement and a consequent reduction in the uncertainty in α to several tenths ppm!

What might the Josephson volt do for voltage standard technology? The present voltage maintenance standard is of course the venerable electrochemical standard cell. It is highly sensitive to temperature and shock and exhibits short-term transients and long-term drifts due to a variety of complex physical and chemical processes. It is unforgiving of clumsy users and it is history-dependent. It is, in short, a rather uncongenial foundation upon which to build a precision measurement system. The Josephson junction, on the other hand, provides a frequency-to-voltage transducer with a transduction ratio which is believed to be simply a fundamental physical constant, independent of all environmental parameters. It thus has all the desirable characteristics of a true atomic or quantum voltage standard. However, the standard cell has been with us so long that its care and feeding is now a highly developed art. It is a tough competitor. Under carefully controlled laboratory conditions, cells can be intercompared with a precision of better than 0.01 ppm. Interlaboratory or international voltage comparisons require physical transport of cells, however, and these are likely to involve uncertainties of several tenths of a ppm or greater. The long-term stability of carefully maintained groups of cells has been shown to be at the one ppm level or better [19]. If the Josephson volt is to supplant the electrochemical volt in practice it must be demonstrated that it can surpass this performance. Although it is not fundamentally essential, it is generally accepted that it must do so on the standard cell's own ground at the 1 V level rather than at the 1–10 mV level at which Josephson junctions normally work. This means that the necessary voltage ratio circuitry and its limitations must be considered as part of the Josephson voltage standard system in comparing performance. There is now evidence [23] (to be discussed below) that a Josephson voltage standard operating at the 1 V level with a precision and long- and short-term stability of better than 0.1 ppm is now realizable. If the current balance and related devices were replaced by the Josephson junction as the defining electrical unit device, an absolute accuracy at this level would be readily accessible.

How far can the method be pushed? The early ac Josephson effect determinations of $2e/h$ mentioned above all used conventional potentiometers with some modification. It has become apparent that further improvements in accuracy can be obtained by using instruments designed specifically for this experiment [20, 24, 23, 25, 26]. In this connection it is important to realize that a Josephson junction is a tunable voltage source. The variable element of a conventional potentiometer is therefore unnecessary. The uncertainties associated with this element can thus be eliminated, at the expense of the flexibility required in the conventional instrument. All that is needed is a highly stable and accurately determined fixed-resistance ratio coupled with a stable current source [23, 24]. This provides two voltages with an accurately known ratio which can be used for simultaneous comparison with a Josephson junction voltage and a standard cell voltage.

Another factor which can contribute to improved accuracy is an increase in the voltage available from the Josephson device. The early determinations involved junction voltages below a millivolt. An increase in this voltage permits a reduction of the required voltage (resistance) ratio and allows determination of this ratio with reduced uncertainty. It also mitigates a variety of other experimental problems which contribute to the uncertainty. The total available voltage can be increased by increasing the voltage available from a single junction and/or by operating several junctions in series. It has been shown theoretically [9] that the amplitude j_1 of the Josephson supercurrent decreases at high voltages but remains fairly large up to voltages several times $2\Delta/e$ (2Δ is the superconducting energy gap). $2\Delta/e$ is typically between 1 and 3 mV. According to the simple theory of the Shapiro steps, the amplitude of the nth step is proportional to $j_1 J_n(2eV_{rf}/h\nu)$, where J_n is the ordinary Bessel function of order n and V_{rf} is the induced rf voltage across the junction at frequency ν [16]. The first (largest) maxima of J_n decrease as $n^{-1/3}$ for large n; together with the slow decrease in j_1, this indicates that the steps should decrease fairly slowly in amplitude with increasing voltage. This is commonly observed in practice. In order to maximize the amplitude of a given step, it is necessary that there be enough rf power available to reach the first maximum of J_n, i.e., $2eV_{rf}/h\nu \sim n$. This can be achieved with several hundred milliwatts for $n \sim 500$ in tunnel junctions by designing the junction so that it has a geometric self-resonance at the desired rf input frequency.

Ultimately the maximum usable single-junction voltage in a $2e/h$ experiment is limited by noise. If the noise is negligible the steps have no detectable voltage width; the experimental upper limits are about 0.1 nV for steps up to 10 mV in tunnel junctions [27], 1 nV at 1.8 mV for point contacts [28], and 10^{-17} V for steps at 1 nV in SNS junctions [29], corresponding to upper limit fractional widths of 10^{-6} to 10^{-8} in these cases. In the presence of noise the steps may be rounded at the upper and lower ends or may become observably nonvertical ($dI/dV < \infty$).

Such nonvertical steps were reported by PLDT, who found that the steps remain symmetrical about the zero noise position (to the nanovolt level), a result consistent with detailed analyses of the effect of noise on steps given by Stephen [30], by Kose and Sullivan [31], and by Clarke, Pippard, and Waldram [32]. Nonvertical steps are therefore still usable in $2e/h$ experiments, but require the exercise of great care, since failure to locate and bias at the precise center of the step can lead to errors. It is obviously far more desirable to work only with vertical steps. Steps have been detected to voltages as high as 17 mV by McDonald et al. in Nb-Nb point contacts [33], but the resolution was not sufficient to permit determining if the steps were vertical at the nanovolt level. Finnegan et al. have obtained vertical steps with amplitudes $\sim 20\ \mu A$ at voltages exceeding 10 mV using Pb-PbO_x-Pb tunnel junctions and a shielded room to exclude external noise [23]. Voltages approaching 10 mV have been obtained under similar conditions by Page et al. [26]. It has been established, therefore, that usable steps can be achieved at voltages near 10 mV in single junctions, and probably at voltages several times larger using several junctions in series. In the latter case, independent biasing of each junction is required. Finnegan et al. also report obtaining voltages exceeding 10 mV with several junctions in series [23]. At present, the inconvenience and complexity which the use of series junctions adds is not justified beyond two or three junctions in series, although voltages beyond 100 mV could probably be achieved in this way if one needed them badly enough.

Preliminary results of a determination of $2e/h$ incorporating these improvements have been reported by Finnegan et al. [23]. The quoted total uncertainty of this result is 0.46 ppm. The statistical uncertainty of the data and uncertainties associated with the voltage comparator circuitry contributed about 0.1 ppm to the total. The remainder was associated with transfer and maintenance of the voltage standard. I believe that this published result and the papers which follow this one indicate that the overall uncertainties of the current generation of $2e/h$ determinations will be controlled by uncertainties in the electrochemical voltage references used, rather than by potentiometer calibration uncertainties as in the previous generation. This means that any significant further improvement in our knowledge of $2e/h$ in terms of an as-maintained volt can come only if the precision and stability of the as-maintained volt itself is improved. It also suggests that it may soon be time to relegate the electrochemical standard cell to the role of a convenient voltage transfer device, and to carry out precision voltage intercomparisons and certain fundamental constant determinations solely in direct contact with Josephson voltage references. It should also be noted that there are operational and conceptual advantages in specifying emfs in terms of Josephson frequencies rather than volts, much as magnetic fields are often specified in terms of proton resonance frequencies. All of these considerations are reflected in the fact that ac Josephson effect determinations of $2e/h$ (or standard cell emfs, depending on your point of view) are under way or are being planned in a large number of national laboratories, including those of the United States, Great Britain, Canada, Australia, Japan, West Germany, France, the Soviet Union, Sweden, and Italy.

Once the transition from standard cell emfs to Josephson frequencies is made, still further improvements in Josephson volt technology may be desirable. In this connection, studies are being made in several laboratories of the possibilities of all cryogenic systems. Voltage comparator resistance networks operating in superfluid liquid helium would function in an insulating fluid having a thermal conductivity several orders of magnitude larger than that of copper at room temperature. Superconducting interconnections and switches can be used. Null detectors based on Josephson junctions can have noise-limited sensitivities orders of magnitude greater than the best room temperature null detectors. Successful exploitation of such advantages in a practical system will require a great deal of research and ingenuity, but attractive possibilities do exist.

Finally, what of the fundamental quantitative reliability of the ac Josephson effect? In his paper, Scalapino indicates that there are now no clear theoretical grounds for expecting any corrections at all to the Josephson frequency-voltage relation. Corrections have been suggested at the 10^{-9} to 10^{-10} level, but these are open to question, and in any case they arc of no concern at the levels of precision of practical interest in the near future. However, if I may paraphrase a current slogan, I would suggest that it is unwise for an experimentalist to trust any theorist, whether he is over or under thirty. Accordingly, I believe that it is incumbent on the experimentalists in the field to continue rigorous testing of the frequency-voltage relation to the limits of practical possibility.

What has already been done? In the sourse of their determination of $2e/h$, PLDT found the frequency voltage ratio to be independent of (1) junction material (Sn, Pb, Nb, Ta, Nb_3Sn); (2) type of junction (tunnel junctions and point contacts); (3) temperature ($0.3 \leq T/Tc \leq 0.9$); (4) magnetic field (0 to 10 G); (5) step number ($n \lesssim 70$); (6) microwave frequency and power (10 GHz and 70 GHz); and (7) whether the ratio was measured using Shapiro steps or radiation emission, all at about the 2 ppm level [15]. Clarke has compared the electrochemical potentials of steps in SNS junctions of different materials, irradiated by the same rf source [29]. The high sensitivity afforded by the use of a slug superconducting galvanometer as a null detector in this differential experiment allowed Clarke to demonstrate that the step potentials for lead, tin, and indium SNS junctions were identical to within 0.01 ppm. In their experiments, Petley and Morris worked near 36 GHz and used slugs rather than tunnel junctions or point contacts, and were thus able to investigate several different Pb-Sn-Cd alloys [21]. The internal agreement of their results and the

external agreement with the result of PLDT further confirms independence of material and frequency at a level of about 2 ppm. Finnegan et al. have extended the frequency range to 891 GHz in a differential experiment in which the potential of a step induced by 9.48 GHz radiation in a Pb-PbO_x-Pb tunnel junction was compared directly with the potential of a step induced by HCN laser radiation in a Nb_3Sn–Nb_3Sn point contact [28]. The result indicated that the frequency-voltage ratio was the same within 1.5 ppm. Recently Finnegan et al. have reinvestigated the temperature, magnetic field, and step number/voltage dependence and found independence to within several parts in 10^8 [25, 27]. An experiment now underway at Stanford is aimed at checking material independence at a level of 10^{-9} or 10^{-10} [34]. Taken together, these results show that the frequency-voltage ratio is independent of a wide variety of experimental parameters at the 10^{-8} level. We may therefore contemplate using the ac Josephson effect in precision measurements at this level, but our confidence should not be so complete as to blind us to the need for continuing rigorous tests.

It should be noted that all of these experiments really test the independence of the *electrochemical potential*, not voltage, so that they provide no information about such questions as possible corrections to the electron charge in metals [35]. On the other hand, it is the electrochemical potential, not the voltage, which is involved in the determination of $2e/h$ and in voltage standard applications. For discussion of this point see PLDT [15] and McCumber [36].

4. Determination of the Flux Quantum and the Electron Compton Wavelength

We come now to a discussion of measurements which depend on the quantum phase coherent properties of superconducting circuits containing Josephson junctions. Consider the device shown in figure 4. It is a thin film superconducting ring on a suitable substrate (e.g., a quartz rod) interrupted by a Josephson junction, shown here as a Dayem bridge. In the superconductor, the gradient of the phase is related to the canonical momentum p of a Cooper pair by the usual quantum relation

$$\hbar\nabla\varphi = \mathbf{p} = 2m_e\mathbf{v} + 2e\mathbf{A}. \tag{4}$$

The phase difference φ across the junction is thus

$$\varphi = \hbar^{-1}\int_1^2 \mathbf{p}\cdot d\mathbf{r} \tag{5}$$

where the path of integration extends around the ring from side 1 of the junction to side 2 of the junction. If we neglect for the moment the small contribution from the junction region, we can close the integration path around the ring and write

$$\varphi = \hbar^{-1}\oint \mathbf{p}\cdot d\mathbf{r} = (2/\hbar)\oint (m_e\mathbf{v} + e\mathbf{A})\cdot d\mathbf{r}. \tag{6}$$

The circulating supercurrent in the junction and the ring is then given by

$$I = I_1 \sin[(2/\hbar)\oint (m_e\mathbf{v} + e\mathbf{A})\cdot d\mathbf{r}]. \tag{7}$$

Detailed analysis shows that this expression remains valid when the contribution of the junction region to the phase integral is properly included. The supercurrent appears implicitly on the right hand side of eq (7) because the pair velocity v is directly related to the pair supercurrent and because the supercurrent contributes to the vector potential A. The response of such a ring to an external magnetic field, for example, can therefore be quite complex [37]. For present purposes it is sufficient to observe that (1) if $\mathbf{A}_x$ is the vector potential due to an external source, $\oint \mathbf{A}_x\cdot d\mathbf{r}$ is the externally applied flux Φ_x linking the ring, and I is periodic in Φ_x with period $h/2e$; (2) if Φ_x is fixed and a pair velocity v is produced by mechanically rotating the ring with angular frequency Ω, I is periodic in Ω with period $h/4m_eS$, where S is the area of the ring. The first effect was first demonstrated experimentally by Jaklevic, Lambe, Silver, and Mercereau [38] and provides an alternative means for determining $h/2e$, which in this incarnation is called the "superconducting flux quantum." The second effect was first demonstrated experimentally by Zimmerman and Mercereau [39] and provides a means for determining h/m_e or, with the addition of a factor c, the electron Compton wavelength $\lambda_C = h/m_ec$.

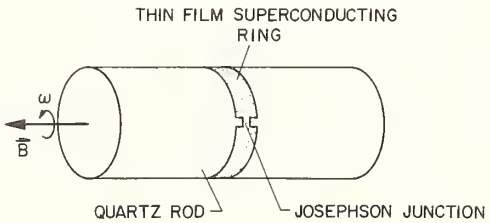


FIGURE 4. *Weak link superconducting ring.*

The most convenient method of monitoring the ring supercurrent is to couple the ring to the coil of a driven radiofrequency resonant circuit [40]. The periodic variations of the supercurrent are reflected in a periodic modulation of the amplitude of the rf voltage across the resonant circuit. This system has recently been analyzed in detail by Simmonds and Parker [41].

The role which such measurements may play in improving our knowledge of the fundamental constants has been discussed by Taylor, Parker, and Langenberg [19]. If the flux quantum experiment is performed in a precision solenoid like those used for low-field γ_p' measurements, and the field determined from the coil geometry and current, it measures $2e/h$. It does so in terms of an as-maintained ampere rather than an as-maintained volt, but otherwise the same considerations apply as for a $2e/h$ determination from the ac Josephson effect. The flux

quantum experiment is considerably more difficult, however, primarily because it requires determination of the effective area of the ring in order to relate field and flux. The size of the ring is ultimately limited by the condition that the thermal flux noise must be less than one flux quantum. This implies that the inductance of the ring, including a contribution due to the kinetic inductance of the supercurrent, must be less than about $h^2/(2e)^2kT$ [41]. So far, the maximum usable ring diameter has been several millimeters, although further refinements of technique may permit use of one centimeter rings [42]. In order to achieve an uncertainty of 10 ppm in the area, the ring diameter must be determined to 5 ppm (100 Å–500 Å) *at liquid helium temperatures*! One suitable substrate material, fused silica, expands by about 75 ppm on cooling to 4 K. Either very careful measurements of the fractional length change or direct measurements of ring diameter at low temperatures will be required. The penetration of the magnetic field into the superconductor must also be taken into account. This makes the effective area of the ring larger than the substrate area by something like 10 ppm. These difficulties will undoubtedly eventually be overcome, and a 10 ppm measurement appears feasible.

A somewhat more interesting version of the flux quantum experiment is one in which the field is determined in terms of the proton precession frequency. Taylor, Parker, and Langenberg [19] have shown that the result of such an experiment can be expressed in the form

$$\lambda_C = \frac{h}{m_e c} = \frac{2\omega_{p0}' S}{c(\mu_p'/\mu_B)}, \qquad (8)$$

where ω_{p0}' is the proton precession frequency (in water) corresponding to the magnetic field increment which causes the flux in the ring to change by exactly one flux quantum. We see that this version amounts to a determination of the electron Compton wavelength and, if we recall that $\alpha^{-2}=2R_\infty\lambda_C$, to a direct determination of the fine structure constant.

The rotation experiment also measures essentially λ_C and therefore also directly determines α. It shares with all ring interferometer experiments the area determination difficulty. It is particularly interesting, however, because it raises and may be able to answer several fundamental questions. Josephson has raised the question whether the electron mass which appears in the ring supercurrent period should be corrected for the relativistic mass equivalent of the electron binding energy (the work function) or some other energy on a scale characteristic of the solid state [43]. For example, one may also ask whether there are any relativistic mass corrections of order $(v_F/c)^2$, where v_F is the Fermi velocity. These would be 10 ppm effects. This experiment is being actively pursued by Simmonds and Parker. A measurement with an uncertainty of 0.1 percent has been reported by Simmonds, Parker, and Nisenoff [44], and further improvements in the accuracy should be forthcoming.

5. Determination of h/M_{He}

The superfluid state of helium (He II) is a macroscopic phase-coherent quantum state like the superconducting state, except that it is uncharged. Effects analogous to the Josephson effects in superconductors are expected in He II. The analog of the induced step or Shapiro step phenomenon has in fact been observed, first by Richards and Anderson [45] and then by Khorana and Chandrasekhar [46]. The "Josephson junction" in these experiments is a small hole (~15 μm diam) connecting two containers of superfluid helium (see fig. 5). The incident radiation is sonic radiation at ~100 kHz supplied by an ultrasonic transducer placed near the orifice. The chemical

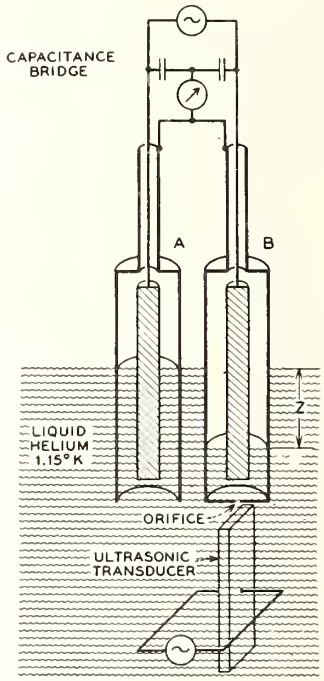


FIGURE 5. *Experimental apparatus for observation of the analog of Shapiro steps in superfluid helium.*
(from P. L. Richards and P. W. Anderson, Phys. Rev. Letters **14**: 540 (1965)).

potential difference between the two containers is $M_{He}gz$, where M_{He} is the mass of the helium atom, g is the gravitational acceleration, and z is the height difference. The height difference is monitored by constructing the two helium containers as coaxial capacitors and connecting them in a capacitance bridge. The bridge unbalance is thus a measure of the capacitance unbalance and this depends on the relative amounts of helium in the two containers. The height difference is found to be quantized (like the Shapiro steps) according to $M_{He}gz=(n/m)h\nu$, where ν is the ultrasonic frequency and n and m are integers. (Such fractional subharmonics are also commonly observed in weak link Shapiro step experiments). An example of these steps is shown in figure 6.

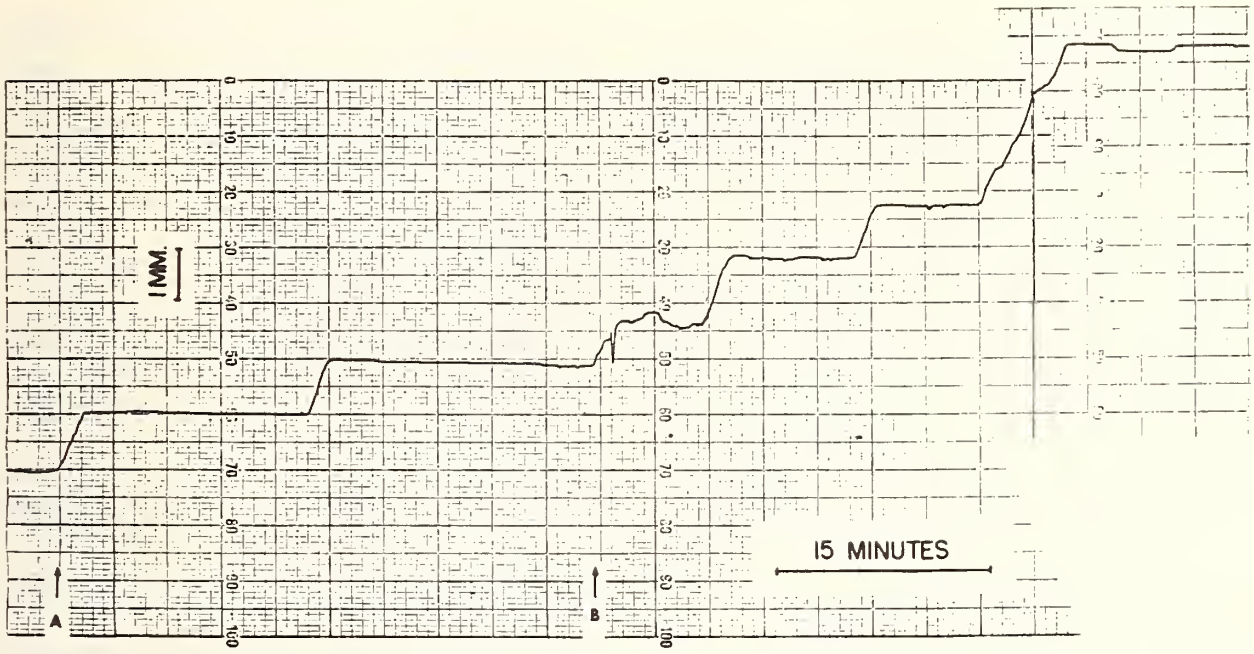


FIGURE 6. *Induced steps in superfluid helium.*
(from B. M. Khorana and B. S. Chandrasekhar, Phys. Rev. Letters **18**, 230 (1967)).

This experiment obviously provides a means for measuring h/M_{He}. Since it gives information on the mass of the helium atom in absolute units, it can be combined with the mass in atomic mass units to provide a value of Avogadro's number. In view of the uncertainties associated with determination of the Faraday, the only present source of information on Avogadro's number, this is of some importance [19]. h/gM_{He} is about 1 cm/MHz, so that for a 10 ppm measurement we are talking about height determinations with optical precision. This is a very difficult experiment, but it is quite likely that it can be developed to a point where it will be of real importance to our knowledge of the fundamental constants. A measurement at the 1 percent level has been reported by Khorana and Douglass [47].

6. Conclusion

I would like to close by stating the obvious: Macroscopic quantum phase coherence and the Josephson effects now play an important role in the field of precision measurement and the fundamental constants. Further significant contributions from this field of solid state physics can be expected in the years to come.

7. Acknowledgments

I would like to take this opportunity to express my thanks to the many colleagues who have contributed to my education on the Josephson effects during the past few years. At the risk of omitting valued friends, I would like to mention specifically J. T. Chen, J. Clarke, A. J. Dahm, A. Denenstein, T. F. Finnegan, B. D. Josephson, J. E. Mercereau, W. H. Parker, D. J. Scalapino, J. R. Schrieffer, and B. N. Taylor.

8. References

[1] Josephson, B. D., Phys. Letters **1**, 251 (1962).
[2] Anderson, P. W., and Dayem, A. H., Phys. Rev. Letters **13**, 195 (1964); Dayem, A. H., and Wiegand, J. J. Phys. Rev. **155**, 419 (1967).
[3] Friehertshauser, P. E., Notarys, H. A., and Mercereau, J. E., Bull. Am. Phys. Soc. **13**, 1670 (1968); Proc. the Intern. Conf. on the Science of Superconductivity (Stanford, 1969), to he published in Physica.
[4] Zimmerman, J. E., and Silver, A. H., Phys. Rev. **141**, 367 (1966).
[5] Clarke, J., J. Phys. **29**, Colloque C2, Suppl. to 2–3 (1968); Proc. Roy. Soc. **A308**, 447 (1969).
[6] Clarke, J., Phil. Mag. **13**, 115 (1966).
[7] For reviews discussing the derivation of the Josephson equations and many of their consequences, see the review by Josephson, B. D., Adv. Phys. **14**, 419 (1965); the review hy Anderson, P. W., in Progress in Low Temperature Physics, Vol. **5**, ed. C. J. Gorter (North-Holland Puhl. Co., Amsterdam, 1967); the articles by Josephson, B. D., and Mercereau, J. E., in Superconductivity, ed. R. D. Parks (M. Dekker, New York, 1969); and the articles hy Scalapino, D. J., Mercereau, J. E., Langenberg, D. N., and Wyatt, A. F. G., in Tunneling Phenomena in Solids, ed. E. Burstein and S. Lundqvist (Plenum Press, Inc., New York, 1969).
[8] Ambegaokar, V., and Baratoff, A., Phys. Rev. Letters **10**, 486 (1963); **11**, 104 (E) (1963).
[9] Riedel, E., Z. Naturforsch. **19A**, 1634 (1964); Werthamer, N. R., Phys. Rev. **147**, 255 (1966); Scalapino, D. J., and Wu, T. M., Phys. Rev. Letters **17**, 315 (1966).
[10] Giaever, I., Phys. Rev. Letters **5**, 147 (1960).

[11] Stewart, W. C., Appl. Phys. Letters **12,** 277 (1968); McCumber, D. E., J. Appl. Phys. **39,** 3113 (1968); Clarke, J., Pippard, A. B., and Waldram, J. R., in Proc. Intern. Conf. on the Science of Superconductivity, Stanford, 1969, to be published in Physica; Waldram, J. R., Pippard, A. B., and Clarke, J., Phil. Trans. Roy. Soc. (London) **A268,** 265 (1970). Buckner, S. A., Chen, J. T., and Langenberg, D. N., Phys. Rev. Letters **25,** 738 (1970).
[12] Swihart, J. C., J. Appl. Phys. **32,** 461 (1961).
[13] Fiske, M. D., Rev. Mod. Phys. **36,** 221 (1964); Coon, D. D., and Fiske, M. D., Phys. Rev. **138,** A744 (1965); Eck, R. E., Scalapino, D. J., and Taylor, B. N., in Proc. Ninth Intern. Conf. on Low Temperature Physics, ed. J. D. Daunt et al. (Plenum Press, Inc., New York, 1965); Chen, J. T., Finnegan, T. F., and Langenberg, D. N., in Proc. Intern. Conf. on the Science of Superconductivity, Stanford, 1969, to be published in Physica.
[14] Yanson, I. K., Svistunov, V. M., and Dmitrenko, I. M., Zh. Eksperim. i Teor. Fiz. **48,** 976 (1965) [Soviet Phys.-JETP **21,** 650 (1965)]; Langenberg, D. N., Scalapino, D. J., Taylor, B. N., and Eck, R. E., Phys. Rev. Letters **15,** 294 (1965); Langenberg, D. N., Scalapino, D. J., and Taylor, B. N., Proc. IEEE **54,** 259 (1966); Dayem, A. H., and Grimes, C. C., Appl. Phys. Letters **9,** 47 (1966); Zimmerman, J. E., Cowan, J. A., and Silver, A. H., Appl. Phys. Letters **9,** 253 (1966).
[15] Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177,** 639 (1969).
[16] Shapiro, S., Phys. Rev. Letters **11,** 80 (1963); Shapiro, S., Janus, A. R., and Holly, S., Rev. Mod. Phys. **36,** 223 (1964).
[17] Langenberg, D. N., Parker, W. H., and Taylor, B. N., Phys. Rev. **150,** 186 (1966).
[18] Parker, W. H., Taylor, B. N., and Langenberg, D. N., Phys. Rev. Letters **18,** 287 (1967); Langenberg, D. N., Parker, W. H., and Taylor, B. N., in Proc. Third Intern. Conf. on Atomic Masses, ed. R. C. Barber (University of Manitoba Press, Winnipeg, Canada, 1968).
[19] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41,** 375 (1969).
[20] Taylor, B. N., Parker, W. H., Langenberg, D. N., and Denenstein, A., Metrologia **3,** 89 (1967).
[21] Petley, B. W., and Morris, K., Phys. Letters **29A,** 289 (1969); Metrologia **6,** 46 (1970).
[22] Denenstein, A., Finnegan, T. F., Langenberg, D. N., Parker, W. H., and Taylor, B. N., Phys. Rev. B **1,** 4500 (1970).
[23] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24,** 738 (1970).
[24] Denenstein, A., Ph.D. Thesis, University of Pennsylvania, 1969.
[25] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Bull. Am. Phys. Soc **15,** 654 (1970).
[26] Page, C. H., Harris, F. K., Fowler, H. A., Toots, J., and Olsen, P. T., Bull. Am. Phys. Soc. **15,** 654 (1970).
[27] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev., in press.
[28] Finnegan, T. F., Denenstein, A., Langenberg, D. N., McMenamin, J. C., Novoseller, D. E., and Cheng, L., Phys. Rev. Letters **23,** 229 (1969).
[29] Clarke, J., Phys. Rev. Letters **21,** 1566 (1968).
[30] Stephen, M. J., Phys. Rev. **186,** 393 (1969).
[31] Kose, V. E., and Sullivan, D. B., J. Appl. Phys. **41,** 169 (1970).
[32] Waldram, J. R., Pippard, A. B., and Clarke, J., Phil. Trans. Roy. Soc. (London) **A268,** 265 (1971).
[33] McDonald, D. G., Kose, V. E., Evenson, K. M., Wells, J. S., and Cupp, J. D., Appl. Phys. Letters **15,** 121 (1969).
[34] Hamilton, W. O., and Fairbank, W. M., private communication.
[35] Nordtvedt, K., Jr., Phys. Rev. **B1,** 81 (1970); Langenberg, D. N., and Schrieffer, J. R., Phys. Rev. **B3,** 1776 (1971); Hartle, J. B., Scalapino, D. J., and Sugar, R. L., Phys. Rev. **B3,** 1778 (1971).
[36] McCumber, D. E., Phys. Rev. Letters **23,** 1228 (1969).
[37] Silver, A. H., and Zimmerman, J. E., Phys. Rev. **157,** 317 (1967).
[38] Jaklevic, R. C., Lambe, J., Silver, A. H., and Mercereau, J. E., Phys. Rev. Letters **12,** 159 (1964); Jaklevic, R. C., Lamhe, J., Mercereau, J. E., and Silver, A. H., Phys. Rev. **140,** A1628 (1965).
[39] Zimmerman, J. E., and Mercereau, J. E., Phys. Rev. Letters **14,** 887 (1965).
[40] Mercereau, J. E., Proc. Symp. on the Physics of Superconducting Devices (University of Virginia, Charlottesville, 1967).
[41] Simmonds, M. B., and Parker, W. H., J. Appl. Phys. **42,** 38 (1971).
[42] Parker, W. H., private communication.
[43] Josephson, B. D., private communication.
[44] Simmonds, M., Parker, W. H., and Nisenoff, M., Bull. Am. Phys. Soc. **13,** 1668 (1968).
[45] Richards, P. L., and Anderson, P. W., Phys. Rev. Letters **14,** 540 (1965).
[46] Khorana, B. M., and Chandrasekhar, B. S., Phys. Rev. Letters **18,** 230 (1967).
[47] Khorana, B. M., and Douglass, D. H., Jr., Bull. Am. Phys. Soc. **14,** 96 (1969).

DISCUSSION

B. N. Taylor: I would like to make one comment concerning the fine structure constant and γ_p. As Mr. Driscoll pointed out yesterday, the measurement of proton precession frequency is well under control, being somewhere in the neighborhood of one part in 10^7. The basic problem with improving γ_p experiments (and therefore a more accurate value of α) is measuring the solenoid dimensions. Anyone who can think of a way to improve the present techniques for measuring solenoid dimensions, or who can build a solenoid which is easier to measure, can make an important contribution.

Noise in a Josephson Junction*

Patrick A. Lee

Physics Department, Yale University, New Haven, Conn. 06520

and

Marlan O. Scully

Physics Department and Materials Science Center, Massachusetts Institute of Technology, and

Physics Department and Optical Science Center, University of Arizona, Tucson, Ariz. 85721

We consider the problem of a Josephson Junction interacting with a quantized electromagnetic field contained in a microwave cavity. Inasmuch as the cavity has a quality factor Q and an external current J flows in and out of the junction, this is a problem in non-equilibrium statistical mechanics. Using a technique that eliminates adiabatically the radiation field directly from the total density matrix equation of motion, we calculate the spectrum of the voltage fluctuations from which the linewidth of the emitted radiation is obtained. As a consequence of this calculation we find that the current-voltage characteristic is slightly modified at finite temperatures. A physical argument for this effect is given and possible experimental observation of this effect is discussed.

Key words: Density matrix; Josephson junction; Josephson radiation; non-equilibrium statistical mechanics; superconductivity.

1. Introduction

It is well known [1, 2] that a superconducting tunnel junction can support a dc current up to a certain critical value without developing any voltage across the junction. However, when an external current in excess of this critical value is forced through the junction, charges pile up on one side of the barrier, and a voltage $V_0=q/C$ is developed, where q is the excess charge on one side of the barrier and C is the capacitance of the junction. In the presence of the voltage, it was predicted that an ac current will oscillate in the junction with a frequency $2eV_0/\hbar$ and radiation will be emitted.

We are interested in calculating the line width of the Josephson radiation. The physical process involved is that when electron pairs tunnel across the junction via interaction with the radiation field, the charge number, and therefore the voltage, fluctuates about the steady-state mean value. This gives rise to a frequency modulation of the radiation and is the main contribution to the line width. Hence we are interested in calculating the voltage correlation function $\langle \delta V(t)\delta V(0)\rangle$, which is simply related to the pair number correlation $\langle \delta N(t)\delta N(0)\rangle$ using the relation $V=2eN/C$. We treat the problem by setting up the equation of motion for the density matrix of the superconductor-radiation system. The voltage fluctuations are calculated by deriving an equation of motion for the reduced density matrix involving the superconductor coordinates alone. This is achieved by noting that the characteristic time of the radiation field $(\nu/Q)^{-1}\approx 10^{-8}$ s is the fastest decay time in the problem. Consequently the radiation field can be eliminated adiabatically. We obtain an equation for the reduced density matrix which is interpreted physically as a flow of probability. The physical origin of the voltage fluctuations is then clearly demonstrated to be shot noise associated with the tunneling of electron pairs. The resulting line-width agrees with that obtained by Stephen [3] using Langevin equations with noise operators. However, in ref. [3] operators are treated as c-numbers and it is felt that the present quantum mechanical treatment is more consistent. Furthermore, as a consequence of the present work, a correction to the usual expression for the current-voltage characteristic is found.

2. Radiation Line Width

The properties of the superconductor relevant to the tunneling problem can be described by three operators S, $S^\dagger$, S_z defined as follows:

$$S=S_L S_R^\dagger, \tag{1}$$

$$S^\dagger=S_R S_L^\dagger, \tag{2}$$

and

$$S_z=\tfrac{1}{2}(S_{z,L}-S_{z,R}) \tag{3}$$

* Work supported by the U.S. Air Force Office of Scientific Research.

where S_L (S_R) destroys a coherent superposition of pairs on the left (right) hand superconductor without disturbing the superconductor from its ground state and $S_{z,L}$ ($S_{z,R}$) describes the number of electron pairs in excess of the value required to maintain charge neutrality. Hence S describes the coherent transfer of electron pairs from left to right, and S_z describes the number of excess electron pairs on the left superconductor. These properties are implied by the following commutation relations:

$$[S, S_z]=S \tag{4}$$

$$[S^\dagger, S_z]=-S^\dagger \tag{5}$$

and for our present purpose we assume that [4]

$$[S, S^\dagger]=0. \tag{6}$$

With this simplification, we note that the operator $SS^\dagger$ commutes with every operator in the system, and can therefore be treated as a c-number. We have chosen a normalization such that

$$S^\dagger S=1. \tag{7}$$

The Hamiltonian for the system consisting of superconductors interacting with the quantized radiation field is given by

$$H=\hbar\Omega a^\dagger a+2e^2S_z^2/C+i\hbar T(a^\dagger S-aS^\dagger) \tag{8}$$

where Ω is the cavity mode frequency and $a^\dagger$ creates a photon in that mode. The first term describes the energy of the free photon, the second term is the energy associated with the capacitance C of the junction and the third term describes the coherent transfer of an electron pair across the barrier with the emission or absorption of a photon. However we must note that photons are constantly leaking out of the cavity which has a quality factor Q and electron pairs are flowing from one side of the barrier to the other as a constant current is forced through the junction. This is a problem in non-equilibrium statistical mechanics and the dynamics cannot be described by the Hamiltonian H alone. Instead we must study the following equation of motion for the density matrix of the superconductor-radiation system:

$$\dot\rho=-(i/\hbar)[H, \rho]+(\dot\rho)_Q+(\dot\rho)_{\text{current}} \tag{9}$$

where

$$(\dot\rho)_Q=-\tfrac{1}{2}(\bar n+1)(\nu/Q)(a^\dagger a\rho+\rho a^\dagger a-2a\rho a^\dagger)$$
$$-\tfrac{1}{2}\bar n(\nu/Q)(aa^\dagger\rho+\rho aa^\dagger-2a^\dagger\rho a) \tag{10}$$

and

$$(\dot\rho)_{\text{current}}=(-A/2)(SS^\dagger\rho+\rho SS^\dagger-2S^\dagger\rho S) \tag{11}$$

Equation (10) describes a cavity at a temperature T so that the average photon number in the absence of other interaetions is given by the Planck function $\bar n=[\exp(\hbar\Omega/kT)-1]^{-1}$. Equation (11) describes a flow of A electron pairs per unit time across the junction. We would like to write out eq (9) in a specific representation. For this purpose we note that the description of the superconductors relevant to our discussion can be spanned by the eigenstates of S_z

$$S_z\,|\,k\rangle=k\,|\,k\rangle. \tag{12}$$

From the commutation relations (4) and (5) and the normalization condition (7), we conclude that

$$S\,|\,k\rangle=|\,k-1\rangle, \tag{13}$$

$$S^\dagger\,|\,k\rangle=|\,k+1\rangle. \tag{14}$$

The state vector $|\,k\rangle$ has the interpretation of representing two superconductors both in their respective B.C.S. states with k excess electron pairs on the left-hand side. The equation of motion for the density operator (eq (9)) can then be written down in n-representation for the field, and pair-number representation for the superconductors. We shall first restrict our attention to the case where $\bar n=0$, i.e., the cavity is at zero temperature.

As discussed in the introduction, we are interested in calculating the pair number correlation function

$$G_N(t_1, t_2)=\langle\delta N(t_1)\,\delta N(t_2)\rangle. \tag{15}$$

This correlation function involves only the superconductor coordinates and hence can be calculated

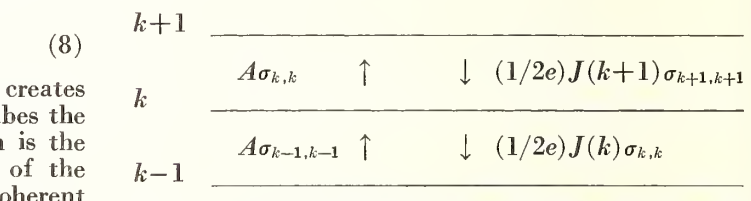


FIGURE 1. *Figure illustrating the flow of probability for having* k *excess pairs on L.H.S. when the cavity temperature* $T=0$. Terms proportional to A are due to external current and drive the system toward higher values of k, while terms resulting from tunneling tend to lower the number of excess pairs, as indicated by the arrows pointing down.

from the equation of motion of the reduced density matrix $\sigma_{k,k'}$ obtained by tracing over the radiation field coordinates as follows:

$$\sigma_{k,k'}=\sum_n \rho_{k,n;k',n}. \tag{16}$$

An equation of motion for the reduced density matrix can be obtained by performing an adiabatic elimination of the photon coordinates from eq (9). This is achieved by noting that the radiation field charaeteristic time $(\nu/Q)^{-1}$ is the fastest decay time in the problem, and hence the radiation field follows the time development of the electron pair population. The time development of the entire system is then described by the time evolution of the reduced density matrix $\sigma_{k,k'}$ alone. The resulting equation of motion for the diagonal element of the reduced density matrix is the following:

$$\dot\sigma_{k,k}=-(1/2e)(J(k)\sigma_{k,k}-J(k+1)\sigma_{k+1,k+1})$$
$$-A(\sigma_{k,k}-\sigma_{k-1,k-1}), \tag{17}$$

where

$$J(k)=\frac{2eT^2(\nu/Q)}{[\Omega-[(2e)^2/\hbar C](k-\frac{1}{2})]^2+(\nu/2Q)^2}. \tag{18}$$

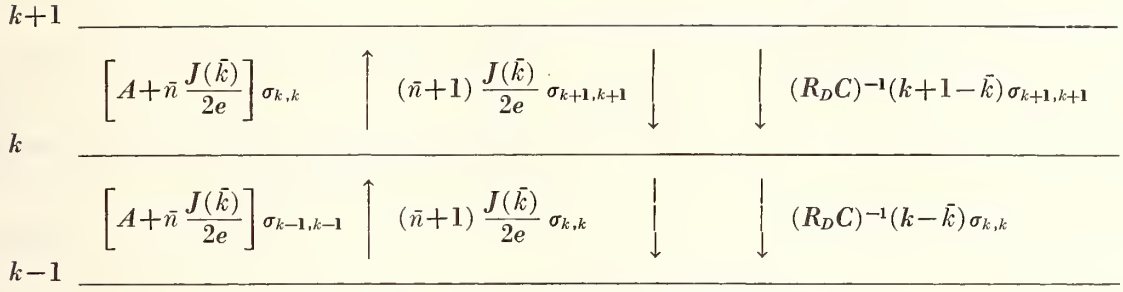


FIGURE 2. *Figure illustrating the flow of probability for having* k *excess pairs on L.H.S. when the cavity temperature is finite, or* $\bar{n}\neq 0$.

Terms proportional to $(A+\bar{n}J(\bar{k})/2e)$ represent the flow of probability toward higher values of k due to external current (A) and "backward" tunneling from right to left due to absorption of thermal photon $(\bar{n})$. Terms proportional to $(\bar{n}+1)J(\bar{k})/2e$ correspond to tunneling of a pair from left to right by stimulated emission of $\bar{n}+1$ photons. Terms proportional to $(R_DC)^{-1}$ tend to restore the average pair number to the mean value $\bar{k}$.

Equation (17) can be interpreted in terms of the flow of probability from the k to $k-1$ and the k to $k+1$ excess pair states at the rate A and $J(k)$, respectively. This is shown in figure 1. Furthermore, detailed balance can be used to obtain the steady state solution for $\sigma_{k,k}$.

Next, let us define $\bar{k}$ as the mean number of pairs at steady state.

$$\bar{k}=\sum_k k(\sigma_{k,k})_{\text{steady state}}. \tag{19}$$

We may approximate eq (18) for $J(k)$ by expanding $J(k)$ about the steady state value $\bar{k}$. Keeping only the first-order correction we have

$$J(k)\approx J(\bar{k})+(k-\bar{k})[\partial J(k)/\partial k]\,|_{k=\bar{k}}. \tag{20}$$

The mean equation of motion for S_z

$$\langle\dot{S}_z\rangle=\sum_k k\dot{\sigma}_{k,k} \tag{21}$$

may now be calculated using eqs (21), (20), and (17). We find

$$\langle\dot{S}_z\rangle=(A-[J(\bar{k})/2e])-(1/2e)(\partial J/\partial\bar{k})(\langle S_z\rangle-\bar{k}). \tag{22}$$

This leads us to identify the average tunneling current with the external current as follows:

$$J_{\text{dc}}=2eA=J(\bar{k}). \tag{23}$$

After making this identification, eq (22) has the following solution:

$$(\langle S_z(t)\rangle-\bar{k})=(\langle S_z(0)\rangle-\bar{k})\times\exp\{-(1/2e)(\partial J/\partial\bar{k})t\}. \tag{24}$$

This shows that $\langle S_z(t)\rangle$ decays to the steady-state value $\bar{k}$ at the rate $(1/2e)(\partial J/\partial\bar{k})$. Noting that the voltage V is given by $V=(2e/C)\bar{k}$, the rate can be written in terms of more physical quantities,

$$(1/2e)(\partial J/\partial\bar{k})=(R_DC)^{-1} \tag{25}$$

where $R_D=(\partial J/\partial V)^{-1}$ is the dynamic resistance of the junction.

The two time correlation function $\langle\delta N(t)\delta N(0)\rangle$ where $\delta N(t)=S_z(t)-\bar{k}$ can be calculated from eq (24) by the regression theorem,

$$\langle\delta N(t)\delta N(0)\rangle=\{\tfrac{1}{2}(J(\bar{k})/2e)R_DC\}\times\exp[-(R_DC)^{-1}t]. \tag{26}$$

Using standard FM theory, eq (26) implies a line width

$$\Delta\nu\,|_{\bar{n}=0}=(8e^3/2\pi\hbar^2)JR_D^2. \tag{27}$$

In the case of finite temperature (nonzero $\bar{n}$), a reduced density matrix equation can also be obtained using a somewhat more complicated technique

$$\dot{\sigma}_{k,k}=-(\bar{n}+1)[J(\bar{k})/2e](\sigma_{k,k}-\sigma_{k+1,k+1})-(\bar{n}[J(\bar{k})/2e]+A)(\sigma_{k,k}-\sigma_{k-1,k-1})-(R_DC)^{-1}[(k-\bar{k})\sigma_{k,k}-(k+1-\bar{k})\sigma_{k+1,k+1}] \tag{28}$$

where

$$J(\bar{k})=\frac{2eT^2(\nu/Q)}{\{\Omega-[(2e)^2/\hbar C][\bar{k}-(\bar{n}+\tfrac{1}{2})]\}^2+(\nu/2Q)^2}. \tag{29}$$

Equation (28) can again be interpreted as a flow of probability as illustrated in figure 2. We note that it is the number of thermal photons $\bar{n}$ and not the total number of photons in the cavity which enters into eq (28) and which is responsible for the shot noise associated with pair tunnelling. To obtain the mean square fluctuation $\langle(\delta N)^2\rangle$ we observe that eq (28) implies that the dc current is made up of a current $(\bar{n}+1)J$ flowing to the right and a current $\bar{n}J$ flowing to the left. Since these currents flow via the transfer of discrete number of electron pairs, we conclude that δN increases or decreases by one at the rate of $(2\bar{n}+1)J/2e$ events per unit time. However, whether δN increases or decreases, $(\delta N)^2=1$. Furthermore, each deviation δN from equilibrium decays like $\exp[-(R_DC)^{-1}t]$. The ensemble average $\langle\delta N^2\rangle$ is given by the number of deviations per unit time,

multiplied by the lifetime of each deviation and the value of $(\delta N)^2$. We then conclude that

$$\langle(\delta N)^2\rangle = (2\bar{n}+1)\,(J/2e)\,\tfrac{1}{2}(R_D C). \tag{30}$$

Hence

$$\langle\delta N(t)\delta N(0)\rangle = \langle[\delta N(0)]^2\rangle \exp[-(R_D C)^{-1}t]$$
$$= \tfrac{1}{2}(2\bar{n}+1)\,(J/2e)R_D C \times \exp[-(R_D C)^{-1}t]. \tag{31}$$

The linewidth can also be calculated and is simply $(2\bar{n}+1)$ times the zero temperature result.

$$\Delta\nu = (2\bar{n}+1)\,(8e^3/2\pi\hbar^2)JR_D{}^2. \tag{32}$$

This linewidth agrees with that obtained by Stephen [3] with the exception that J is given by a slightly modified expression eq (29). In the ease of a tunnel junction the resonance cavity is the oxide layer itself. A finite cavity temperature implies a finite superconductor temperature and the contribution of the quasi-particle current to the voltage fluctuations will have to be included. This contribution has been considered by Scalapino [5] and a combination of the two contributions have been found to be in fairly good agreement with experiments [6].

It is of interest to note that eq (29) for the current $J(\bar{k})$ is different from the zero temperature expression (eq (27)) in that it is a Lorentzian shifted by the amount $[(2e)^2/\hbar C](\bar{n}+\frac{1}{2})$. In the following section we shall give a physical interpretation of the effect.

3. Physical Discussion

We note that the current J can be calculated from eq (9) as follows,

$$J = -2e(\partial\langle S_z\rangle/\partial t) = 2eT(\langle S^\dagger a\rangle + \langle S a^\dagger\rangle). \tag{33}$$

The mean equation of motion for $S^\dagger a$ can be similarly obtained

$$\frac{\partial}{\partial t}\langle S^\dagger a\rangle = -(i\Omega+\nu/2Q)\langle S^\dagger a\rangle + i\frac{2e^2}{\hbar C} \times \langle S_z S^\dagger a + S^\dagger a S_z)\rangle - T\langle S^\dagger S\rangle$$
$$= -\left[i\left(\Omega - \frac{(2e)^2}{\hbar C}\bar{k}\right)+\nu/2Q\right]\langle S^\dagger a\rangle + i\frac{2e^2}{\hbar C}[\langle(S_z-\bar{k})S^\dagger a\rangle + \langle S^\dagger a(S_z-\bar{k})\rangle] - T\langle S^\dagger S\rangle, \tag{34}$$

where $\bar{k}$ is the mean number of pairs at steady state. Usually $\langle(S_z-\bar{k})S^\dagger a\rangle$ and $\langle S^\dagger a(S_z-\bar{k})\rangle$ are assumed to factorize and hence vanish. One then solves eq (34) for the steady-state value of $\langle S^\dagger a\rangle$ and obtains the usual expression (eq (23)) for J from eq (33). However, as we shall show, this factorization cannot be made at finite temperatures, and in fact

$$\langle S^\dagger a(S_z-\bar{k})\rangle + \langle(S_z-\bar{k})S^\dagger a\rangle \approx -(2\bar{n}+1)\langle S^\dagger a\rangle. \tag{35}$$

Putting eq (35) into eq (34) to solve for the steady-state value of $\langle S^\dagger a\rangle$, we obtain

$$\langle S^\dagger a\rangle = \frac{T\langle S^\dagger S\rangle}{i[\Omega - [(2e)^2/\hbar C](\bar{k}-(\bar{n}+\frac{1}{2}))]+\nu/2Q}. \tag{36}$$

Putting eq (36) and its complex conjugate into eq (33) we obtain eq (29) for J.

Before we proceed to analyze this expression for J any further, let us try to understand eq (35) physically. We find it instructive to examine the product $\langle(S_z-\bar{k})a^\dagger a\rangle$. Introducing

$$\delta N = S_z - \bar{k} \tag{37}$$

and

$$\delta n = a^\dagger a - \langle a^\dagger a\rangle, \tag{38}$$

we can write

$$\langle(S_z-\bar{k})a^\dagger a\rangle = \langle\delta N\delta n\rangle. \tag{39}$$

Equation (39) describes the product of the fluctuations from steady state of the number of electron pairs and the number of photons in the cavity. The mean value $\langle\delta N\delta n\rangle$ will be nonvanishing if the fluctuations in the pair number and the photon number are not independent. In fact, they are strongly correlated in the sense that whenever δN decreases by one (an electron pair tunnels from left to right) a photon is emitted into the cavity and δn will increase by one. Similarly when δN increases by one a photon is absorbed and δn decreases by one. Hence, whenever an emission or absorption of a photon takes place the product $\delta N\delta n$ takes on the value negative unity. The ensemble average $\langle\delta N\delta n\rangle$ can be obtained in the same way that $\langle(\delta N)^2\rangle$ was obtained in the last section, except that the photon lifetime $(\nu/Q)^{-1}$ should be used.

$$\langle\delta N\delta n\rangle \approx -(2\bar{n}+1)\,(J/2e)\,(\nu/Q)^{-1}. \tag{40}$$

It is then clear that the nonvanishing of the product $\langle\delta N\delta n\rangle$ is a consequence of the discrete nature of the pair current and the finite lifetime of the electron pair-photon correlation.

Let us note that

$$J = 2e(\nu/Q)\langle a^\dagger a\rangle \tag{41}$$

which simply expresses the fact that the rate at which photons leak out of the cavity is the same as the rate at which electron pairs tunnel across the junction. Using eqs (39), (40), and (41), we conclude that

$$\langle(S_z-\bar{k})a^\dagger a\rangle \approx -(2\bar{n}+1)\langle a^\dagger a\rangle. \tag{42}$$

We now recall that the radiation field a is driven by the force S, and hence its equilibrium value is proportional to S. Replacing a by S on both sides in

eq (42), we obtain

$$\langle (S_z - \bar{k}) a^\dagger S \rangle \approx -(2\bar{n}+1)\langle a^\dagger S \rangle, \qquad (43)$$

which, up to a factor of one half, gives us an understanding of eq (35).

Returning to analyze the implications of eq (29), we find it instructive to rewrite it as follows:

$$J = \frac{2eT^2(\nu/Q)}{\{[\Omega + [(2e)^2/\hbar C](\bar{n}+\frac{1}{2})] - \omega\}^2 + (\nu/2Q)^2} \qquad (44)$$

where

$$\omega = 2eV/\hbar = [(2e)^2/\hbar C]\bar{k}.$$

Equation (44) describes a Lorentzian centered at the frequency $\Omega + [2e^2/\hbar C](\bar{n}+\frac{1}{2})$ as the voltage, or ω, is varied. Experimentally one observes the rising half of the Lorentzian as a step in the I–V characteristic as one increases the voltage V. Equation (44) then predicts a shift in the position of the step towards higher voltage as the temperature, hence $\bar{n}$, is increased. For $V = 10^{-6}$ V, $C = 0.1$ μF, and $T = 1$ K, the shift is of order 10^{-10} V. One possible way of observing this shift experimentally is to expose the resonance cavity to a thermal radiation source. The effective $\bar{n}$ will be enhanced and a larger shift produced.

4. References

[1] Josephson, B. D., Phys. Letters **1**, 251 (1962).
[2] Anderson, P. W., and Rowell, J. M., Phys. Rev. Letters **10**, 230 (1963).
[3] Stephen, M. J., Phys. Rev. Letters **21**, 1629 (1968); and Phys. Rev. **182**, 531 (1969).
[4] For a discussion of the consequences of relaxing this assumption, see Scully, M. O., and Lee, P. A., Phys. Rev. Letters **22**, 23 (1969); Lee, P. A., and Scully, M. O., Ann. N.Y. Acad. Sci. **618**, 387 (1970).
[5] Scalapino, D. J., Proc. Symp. on the Physics of Superconducting Devices, University of Virginia (1967) G-1.
[6] Dahm, A. J., Denenstein, A., Langenberg, D. N., Parker, W. H., Rogovin, R., and Scalapino, D. J., Phys. Rev. Letters **22**, 1416 (1969).

DISCUSSION

D. J. Scalapino: I wonder if you would comment on what chemical potential difference would be across this junction with the sort of shift in frequency that you have shown could occur here?

P. Lee: This is not a shift in frequency. This is just a shift in the position of the step in the I–V curve.

D. J. Scalapino: Maybe I misunderstood. Would you conclude that the Josephson frequency-voltage relation remains valid?

P. Lee: Yes.

Pull-in Effect of Josephson Oscillators and an Improved Method of e/h Determination

V. Kose, H. Fack, F. Melchert, and H.-J. Schrader

Physikalisch-Technische Bundesanstalt, Braunschweig, West Germany

Constant voltage steps in dc current-voltage characteristics occur by assuming pull-in phenomena of Josephson oscillators. It is demonstrated that the step amplitude depends on the frequency and the amplitude of the external applied signal. A much more accurate measurement system for e/h determination than those previously published is described.

Key words: Fundamental physical constants; Josephson effect; superconductivity; $2e/h$.

1. Josephson Oscillators

A Josephson junction functions as an oscillator when a dc voltage $V \neq 0$ is applied to the junction. The supercurrent oscillates sinusoidally with a frequency

$$f_j = (2e/h)\,V \tag{1}$$

where e is the electron charge and h is Planck's constant. When, however, a current source instead of a voltage source [1] is applied and therefore a dc current I is fed into the weak link, the supercurrent turns out to be nonsinusoidal. The fundamental frequency of the oscillation is given by eq (1) where V now means the time averaged ac voltage of the oscillator.

All external electrical sources experimentally used in order to supply Josephson junctions can be considered as electrical energy sources with an internal resistance R_i where $0 < R_i < \infty$. It is shown in reference [2] that the differential equation describing the electrical behavior of the junction in conjunction with these sources is the same as if the junction had been supplied by a current source. Only the values for the coefficients are different. So without lack of generality we confine ourselves to the discussion of the current source case, although this cannot be verified experimentally. Note that the situation is quite different for voltage sources because the obtained differential equation is only restricted to this special case $R_i = 0$. However, no experiment is known to which these analytical solutions could be applied. See reference [2] for a rigorous calculation of the differential equations where the transition states and an explanation of the various results are given.

By feeding a dc current I into a point contact junction, the total current through the contact region consists partially of a supercurrent $I_0 \sin\phi(t)$ and of a quasiparticle current $v(t)/R_0$. R_0 is the contact resistance in the normal state and where $v(t) = (\hbar/2e)[d\phi(t)/dt]$ is the time-dependent voltage across the junction. By following Aslamazov, Larkin [3], and McCumber [4] one obtains

$$I = I_0 \sin\phi(t) + (\hbar/2eR_0)[d\phi(t)/dt] \tag{2}$$

where I_0 is the maximum supercurrent at zero voltage and $\phi(t)$ is the time dependent quantum-mechanical phase difference. $\hbar$ means Planck's constant h divided by 2π. The stationary solution of eq (2) is according to reference [3]

$$\frac{v(t)}{I_0 R_0} = \frac{\alpha^2 - 1}{\alpha + \cos(2\pi f_j t)} \tag{3}$$

for $\alpha = I/I_0 > 1$. The time average of $v(t)$ is equal to $V = R_0 I_0 (\alpha^2 - 1)^{1/2}$ and $f_j(\alpha) = (2e/h) R_0 I_0 (\alpha^2 - 1)^{1/2}$, so that eq (1) is satisfied as mentioned above.

The supercurrent—here written as a normalized quantity $\xi(t) = \sin\phi(t)$ is generally non-sinusoidal and is given by

$$\xi(t) = \alpha - [v(t)/I_0 R_0]. \tag{4}$$

2. Pull-in Process of Josephson Oscillators

The Josephson oscillator, as any oscillator, necessarily behaves nonlinearly because it is limited in its amplitude. If such a system is subjected to an external signal or radiation of frequency f which can be very different from the eigenfrequency f_j of the oscillator, one obtains under certain conditions periodic oscillations with $1/f$ instead of $1/f_j$. For selected values of coupling—here given by the amplitude $\hat{I}$ of the external applied ac current $\hat{I}\cos 2\pi f t$—and for a range of frequencies f_j—here $f_j(I)$ depending on the dc current I—the oscillator "forgets" its eigenfrequency and is periodic with $1/f$. This phenomenon is known as "pull-in" and can quite generally be observed in any given oscillator, e.g., a mechanical or an electrical one [5]. The pull-in effect in conjunction with Josephson oscillators was seen at first by Anderson [6].

To study this process we have to extend eq (2), which describes the oscillator by $\hat{I}\cos 2\pi ft$ and find

$$I+\hat{I}\cos 2\pi ft = I_0 \sin\phi(t) + (\hbar/2eR_0)[d\phi(t)/dt]. \quad (5)$$

There is generally no analytical solution of this differential equation except of the time average of the ac voltage $v(t) = (\hbar/2e)[d\phi(t)/dt]$ under the condition of pull-in. As mentioned above, pull-in means that the pair current and the junction voltage have to be periodic with $1/f$. The dc voltage turns out to be constant as long as $v(t)$ and $\sin\phi(t)$ is periodic with $1/f$. The mean value of the time dependent voltage is defined as

$$\langle v(t)\rangle = T^{-1}\int_0^T v(t)\,dt,$$

where T is the period of the oscillation. By using $v(t) = (\hbar/2e)[d\phi(t)/dt]$ we obtain with $T=1/f$

$$\langle v(t)\rangle = (\hbar f/2e)[\phi]_0^T. \quad (6)$$

Because of supposed periodicity of $\sin\phi(t)$, the phase difference ϕ in one period T can only increase by an integral number multiple n of 2π. The number n has to be constant for every pull-in range, so that

$$\langle v(t)\rangle = n(hf/2e); \qquad n=0, 1, 2, \ldots \quad (7)$$

"Substeps" occur in contrast to the voltage source case when $\xi(t) = \sin\phi(t)$ has a shortest period of m/f. A similar procedure as carried out above yields a generalization of eq (7)

$$\langle v(t)\rangle = (n/m)(hf/2e); \qquad n=0, 1, 2, \ldots; m=1, 2, \ldots \quad (8)$$

See also reference [7] for a detailed discussion of the "substeps."

3. Computer Results and Discussion

In order to find the pull-in ranges or, what is equivalent, the current values I for which eq (8) is satisfied, one has to solve eq (5). There are already three solutions of this problem [3, 7, 8]. However we want to consider another aspect of this equation, mainly the frequency dependence of the induced steps.

In figure 1 we plot the I–V characteristic in normalized coordinates $\alpha = I/I_0$ and $\langle v(t)\rangle/(I_0R_0)$ respectively. The parameters $\hat{I}=3I_0$ and $f=(e/h)I_0R_0$ were kept constant, this display. The time dependence of the normalized supercurrent confirms the concept of the pull-in effect where only on a step (see crosspoints in the graph for e.g., $\alpha=1.3$ and $\alpha=1.7$) $\xi(t)$ is *periodic* with $T=1/f$. Circle points symbolize *nonperiodicity* of $\xi(t)$ (e.g., for $\alpha=1.2$). The dc supercurrent shown in the graph as a dashed horizontal line is simply related to the dc voltage by $\bar{\xi}=\alpha-\langle v(t)\rangle/(I_0R_0)$.

It was necessary to have a proof of the constancy of $\langle v(t)\rangle$ in the form of an analytical expression because numerical results are only known to a certain degree of accuracy. This is important because steps of constant dc voltages are used in high precision determinations of e/h based on eq (7) or eq (8). Slanted steps which arise in the presence of internal or external noise will not be considered here [9, 10].

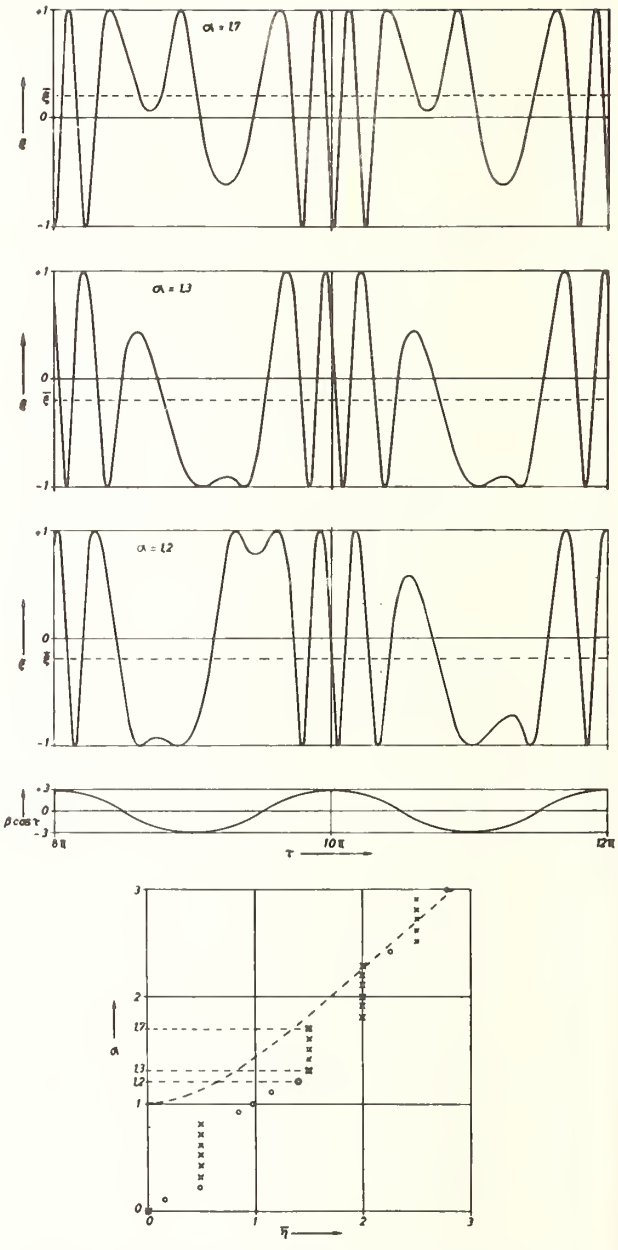


FIGURE 1. *The normalized supercurrent* $\xi(t)=\sin\phi(t)$ *is plotted versus* $\tau=2\pi ft$ *for three different* $\alpha=I/I_0$ *values which can be taken from the current-voltage characteristic below* ($\bar{\eta}=\langle v(t)\rangle/(R_0I_0)$).
For comparison the sinusoidal external applied signal $\hat{I}\cos 2\pi ft$ divided by I_0, i.e., $\beta\cos\tau$, is shown.

Contrary to the voltage source case the differential equation discussed here yields an entirely different solution which is demonstrated in figure 2. By varying the so far fixed parameters, namely the amplitude $\hat{I}$ and the frequency f of the external applied signal, one realizes the disagreement of the results when the

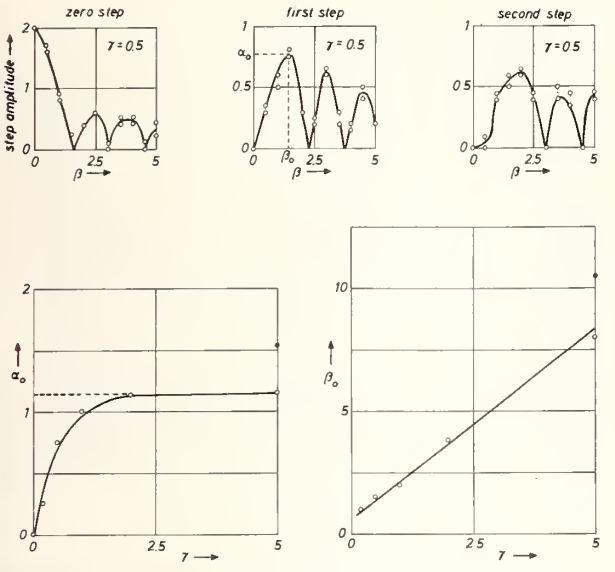


FIGURE 2. *The normalized amplitude of the critical current and the next two steps are plotted versus* $\beta = \hat{I}/I_0$ *for* $\gamma = hf/(2eR_0I_0) = 0.5$.

The first maximum of the first step (α_0) and the corresponding value β_0 is shown as function of γ. (Dark points mean values obtained by including the gap behavior).

junction is supplied by a voltage source instead of a current source. It is obvious by inspection of the graph that the $\hat{I}$ dependence of the step amplitudes does not generally follow Bessel functions which is especially seen in the frequency dependence of the step amplitudes.

The first maximum of the first step (α_0) is plotted versus the normalized frequency and increases significantly for frequencies $\gamma = hf/(2eR_0I_0) < 2$, whereas the voltage source solution does not show any frequency dependence as indicated in the graph by the dashed horizontal line. It seems to be useful for some applications to know the corresponding normalized amplitude $\beta_0 = \hat{I}/I_0$ of the external imposed ac current which is necessary to induce these first maxima of the first step (see fig. 2).

We conclude that constant voltage steps in I–V-characteristics are pull-in ranges of the Josephson oscillator, i.e., that the pair current and the junction voltage are periodic with $1/f$ of the external applied signal. We demonstrated that there is a considerable difference in the results when instead of a voltage source (internal resistance $R_i = 0$) an electrical energy source with $R_i > 0$ is connected with a Josephson junction. It is shown that the step amplitudes do not follow generally simple Bessel functions and the step maxima are not independent of frequency. Earlier discrepancies [11] between theoretical and experimental results should now be overcome by the given analysis.

4. An Improved Method of e/h Determination

Parker et al. [12] have experimentally shown that the steplike current-voltage characteristic of a Josephson junction—as demonstrated in figure 1—can be used to determine the fundamental physical constant e/h with high accuracy. According to the Josephson relation of eq (7) the dc voltage on the nth step and the frequency of the external applied signal have to be measured. The measurement uncertainty in the e/h determination is limited by the voltage-measuring system. There are two methods published by now [12, 13] to determine e/h by the ac Josephson effect. They have one main feature in common: The frequency of the irradiated microwaves was kept constant and the whole comparison of the Josephson voltage with the emf of a standard cell was done by adjusting a seven-decade resistor in conjunction with a voltbox. This resistor was a Kelvin-Varley-divider or a Stabaumatic potentiometer and a voltbox was needed in order to establish a fixed resistance ratio.

The major difference between the system used so far and the arrangement developed at the Physikalisch-Technische Bundesanstalt is that only a *coarse* adjustment of the voltage comparison is done by establishing a voltage ratio with fixed resistors. By taking advantage of the unique property of the Josephson equation that the frequency is proportional to the junction dc voltage, the *fine* adjustment is achieved by varying the frequency of the microwave with a nine-decade frequency synthesizer. The main advantage of this system is the omission of the seven-decade resistor which yields a significant contribution to the overall error of the e/h determination. More details can be taken from figure 3, where the measurement system is shown schematically.

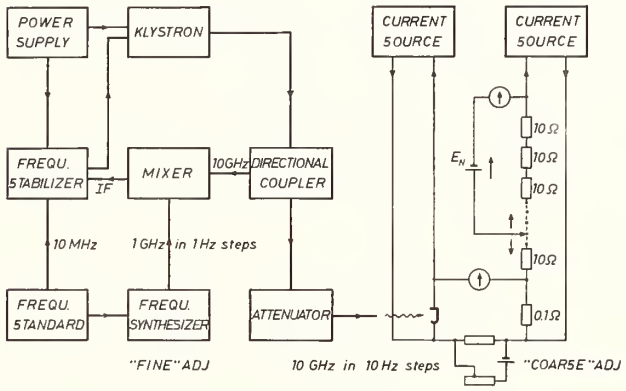


FIGURE 3. *Schematic drawing of the e/h measurement system developed at the PTB.*

By ignoring the synthesizer, the block diagram in figure 3 represents the usual microwave stabilization circuit in order to stabilize an X-band frequency to about 1 part in 10^8. The frequency synthesizer provides a frequency selection at 1 GHz in 1 Hz steps which enables us to adjust the Josephson dc voltage in digits of 1 part in 10^8 via a digital variation of the klystron frequency.

To apply the described system for the e/h measure-

ment, one requirement has to be fulfilled by the Josephson junction. The junction has to be as aperiodic as possible with respect to the absorbed power in the tuned frequency range. This is particularly the case when point contact junctions are used. Parker et al. [12] reported that especially these junctions show "nonvertical" steps. However, Stephen [9], Kose et al. [10], and Finnegan et al. [14] suggest that this is not an intrinsic characteristic of the junctions. Whenever the junction was properly shielded against electrical stray pickup, we could not observe any change in the voltage (<0.1 nV) as the current was varied over the full height of the step. Since point contact junctions do not show the apparent disadvantage of slanted steps, they are as appropriate as tunnel junctions for a high-precision e/h measurement.

The main preferences of the described method of measuring e/h in comparison to the former methods [12, 13] are scheduled below:

1. The *coarse* voltage comparison is done by establishing a fixed resistance ratio (voltbox). The necessary *fine* adjustment is accomplished by means of a frequency synthesizer which does not contribute any further error to the measurement uncertainty. Consequently the voltage-measurement circuit includes no seven-decade-resistor (Kelvin-Varley-divider or Stabaumatic potentiometer) or another resistor with switches and dial settings.
2. The voltage drop across the upper resistor of the voltbox is directly compared with the emf of a standard cell E_N of the primary group of the PTB which is indicated in figure 3. No standard cell transfer is therefore necessary.
3. The voltbox for establishing the voltage ratio V/E_N is composed of two four-terminal resistors. The lower resistor (a series parallel network) in the Josephson circuit has a value of 0.1 Ω and the upper consists of several 10 Ω resistors. The voltage ratio can be measured in itself by means of a potentiometer. A lead resistance correction is avoided.
4. The influences of selfheating, ground loop currents, and the temperature drift of the voltbox are very small.
5. In order to reduce the uncertainty of thermal emf's a Lindeck-circuit is added which eliminates the thermal emf at the beginning of a measuring cycle.
6. The overall time used for measuring e/h including all calibrations is significantly reduced.

Finally it can be stated that an assessment of the experimental errors shows that more than 50 percent of the error items listed in reference [12] and reference [13] are completely avoided by our measurement system whereas the other errors are substantially reduced. Since our e/h measurement is still underway, the complete set of random errors is not yet available but it looks very promising to determine e/h with an uncertainty of less than 1 ppm.

NOTE ADDED IN PROOF. Finnegan et al. [15] have also made use of the unique properties of the Josephson junction as a frequency tunable device to improve their measurement accuracy. See also Harvey et al. [16] and Kose et al. [17].

5. References

[1] A voltage source is defined here as an ac or dc source with internal resistance $R_i=0$; a current source is one with $R_i\to\infty$.
[2] Fack, H., Kose, V., and Schrader, H.-J., Messtechnik **79**, 31 (1971).
[3] Aslamazov, L. G., and Larkin, A. I., JETP Letters **9**, 87 (1969).
[4] McCumber, D. E., J. Appl. Phys. **39**, 3113 (1968).
[5] Fack, H., Frequenz **6**, 141 (1952).
[6] Anderson, P. W., Rev. Mod. Phys. **38**, 298 (1966); Anderson, P. W., in Progress in Low Temperature Physics V, p. 1 (ed. C. J. Gorter) (1967).
[7] Sullivan, D. B., Peterson, R. L., Kose, V. E., and Zimmerman, J. E., J. Appl. Phys. **41**, 4865 (1970).
[8] Clarke, J., Pippard, A. B., and Waldram, J. R., paper delivered at International Conference on the Science of Superconductivity, Stanford, California, Aug. 1969, to be published in Physica.
[9] Stephen, M. J., Phys. Rev. **182**, 531 (1969).
[10] Kose, V. E., and Sullivan, D. B., J. Appl. Phys. **41**, 169 (1970).
[11] Grimes, C. C., and Shapiro, S., Phys. Rev. **169**, 397 (1968).
[12] Taylor, B. N., Parker, W. H., Langenberg, D. N., and Denenstein, A., Metrologia **3**, 89 (1967); Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177**, 639 (1969).
[13] Petley, B. W., and Morris, K., Phys. Letters **29A**, 289 (1969); Metrologia **6**, 46 (1970).
[14] Finnegan, T. F., Denenstein, A., Langenberg, D. N., McMenamin, J. C., Novoseller, D. E., and Cheng, L., Phys. Rev. Letters **23**, 229 (1969).
[15] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).
[16] Harvey, I. K., Macfarlane, J. C., and Frenkel, R. B., Phys. Rev. Letters **25**, 853 (1970).
[17] Kose, V., Melchert, F., Fack, H., and Schrader, H. J., PTB-Mitt. **81**, 8 (1971).

DISCUSSION

W. H. PARKER: I assume since you quoted no number it's too early at this stage to venture a guess as to what your final answer will be?

V. KOSE: Right now I can't tell you, because the uncertainty of the voltage ratio by measuring the voltage box is not available yet. Maybe—but this is only a guess—it will be about 0.6 ppm. So I can't tell you anything above the overall uncertainty of our measurement. It also depends on the highest voltage we will obtain from our point contact junctions.

B. N. TAYLOR: What voltage levels are you actually running at—typically?

V. KOSE: Typically we are now at 1 to 2 millivolts, but I think we can get higher. One possibility to increase steps might be the use of higher frequencies as was indicated in the first part of this talk.

Referencing of the U.S. National Volt Against a Josephson Frequency-to-Voltage Source

H. A. Fowler, T. J. Witt, J. Toots, P. T. Olsen, and W. Eicke

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Intercomparison between the U.S. National Volt group of reference cells, and Josephson voltages at the 2 to 10 millivolt level is reported. High accuracy and precision are achieved through a fixed-element Hamon-pair potentiometer, with amalgam-link switching under stirred oil.

Key words: Josephson effect; potentiometer; voltage standard.

We are reporting on the progress of a program to reference the U.S. National Volt group of standard cells to a Josephson frequency-to-voltage source. The experiment contains the following original features:

(1) The working standard cells have been repeatedly compared with the National Reference Group, over an extended period of time.

(2) DC voltage steps of 2 to 10 millivolts are produced by resonant rf excitation of Josephson junctions of the crossed-strip type, of Pb-PbO-Pb composition. Voltage level is chosen for optimum stability and signal-to-noise of the Josephson steps; and the potentiometer voltage level is adjusted accordingly, by current division in a Kelvin-Varley divider.

(3) The 1:100 stepdown ratio from the standard cells to the 10 mV level is achieved with a transposable pair of Hamon networks, each having a chain of ten 100-ohm resistors. Series-to-parallel switching is achieved through mercury-wetted amalgam contacts under stirred oil.

The potentiometer is shown schematically in figure 1. Major elements are the accurate 1:100 resistance ratio between the parallel and series Hamon networks, the constant-current source, and the Kelvin-Varley current divider.

The accurate parallel-to-series ratio is achieved through careful maintenance of resistance ratios among the current and voltage leads of the Hamon networks. Paralleling connections are made through copper links with mercury-wetted amalgam surfaces; these give contact resistances between 0.1 and 1 $\mu\Omega$, which have negligible effect on the network balance. The Hamon networks are matched within one ppm; by rapid, repeated transposition during the course of the measurement, the errors in the parallel-to-series ratio are believed to be eliminated to an accuracy of some parts in 10^8 [1].

The constant-current source is stabilized against a floating mercury battery by a high-gain, low-noise operational amplifier. During a measurement cycle the current drift is of order 0.1 ppm.

The six dial 10 kΩ Kelvin-Varley divider is used to adjust current passing into the 10-Ω Hamon network serving as a Lindeck element. Its inclusion permits operating the Josephson junction at whatever appear

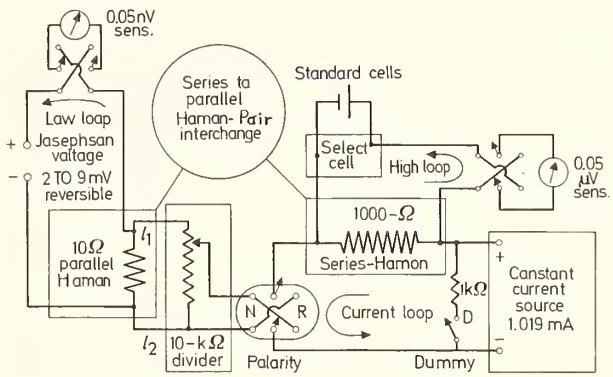


FIGURE 1. *Schematic Circuit of Hamon-Pair Potentiometer.* The one-milliampere current from the constant-current source is adjusted to null the high-loop current against the standard cell. Low-loop balance is produced by selecting the nearest dial setting on the Kelvin-Varley divider, and achieving vernier balance through variation of the Josephson voltage via variation of the applied microwave frequency.

to be its best current steps, usually between 2 and 10 mV. This divider is periodically calibrated against a second stable and accurately known seven-dial divider.

The entire potentiometer is immersed in a large, smoothly-stirred oil bath, to minimize the effects of thermal gradients.

The instrument is described in detail in a forthcoming article [2].

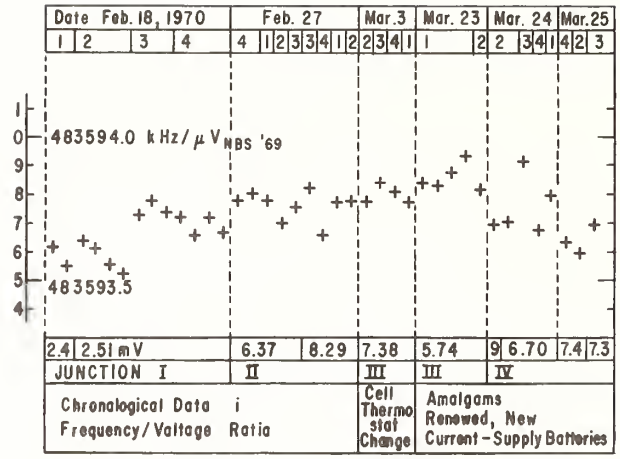


FIGURE 2. *Chronological Data Plot: Josephson Frequency-to-voltage ratio from six different Josephson junctions, measured against the U.S. National Volt Group.*

Dates, Feb. 18–June 23, 1970 inclusive. Working-cell identity for the individual measurement is indicated in the second line. The points are averages of the Hamon-pair transpositions. Thermal emf's in the low-voltage measuring loop are compensated by polarity reversal. The line below the data points shows the Josephson voltage used for each particular set of measurements. Corrections have been applied for divider calibration, lead compensation, working-cell temperature, and offset of the frequency-counter reference crystal.

The complete measurement cycle commences with balancing against the standard cell by adjusting the current-supply output. The Kelvin-Varley divider is then adjusted for an approximate balance against the voltage of the chosen Josephson step. A fine adjustment of the microwave source frequency completes the balance. The polarity of the bias current supplied to the Josephson junction is then reversed, and a balance is made against this negative step with the potentiometer current also reversed. Algebraic addition of the voltage-balances subtracts out the thermal emf in the leads to the junction. One or two more polarity reversals are made, and the current balance against the standard cell, on the series Hamon network, is checked. Following this the Hamon networks are transposed, and the measurement cycle is repeated in the new configuration. Averaging the two results from the transposition symmetrically, eliminates the errors from the series-to-parallel network balance to parts in 10^8.

The four standard cells used with the potentiometer are enclosed in a temperature-regulated air bath, and assigned against the National Reference Group at least once a week.

Figures 2 and 3 are a chronological plot of individual determinations of the frequency-to-voltage ratio, averaged for the Hamon-network-transposition. The upper line, below the observation date, indicates the particular standard cell of our working group, used in that determination. Below the data we list the voltage level at which the measurement was made, the junction on which the measurement was performed, and finally some modifications to the measuring instruments.

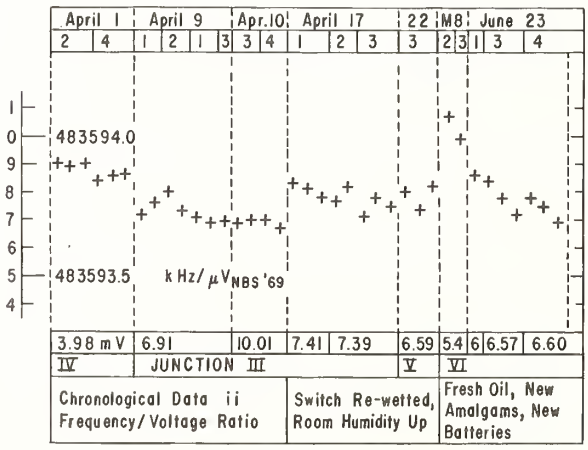


FIGURE 3. *Chronological Data Plot: Josephson Frequency-to-voltage ratio from six different Josephson junctions, measured against the U.S. National Volt Group.*

Dates, Feb. 18–June 23, 1970 inclusive. Working-cell identity for the individual measurement is indicated in the second line. The points are averages of the Hamon-pair transpositions. Thermal emf's in the low-voltage measuring loop are compensated by polarity reversal. The line below the data points shows the Josephson voltage used for each particular set of measurements. Corrections have been applied for divider calibration, lead compensation, working-cell temperature, and offset of the frequency-counter reference crystal.

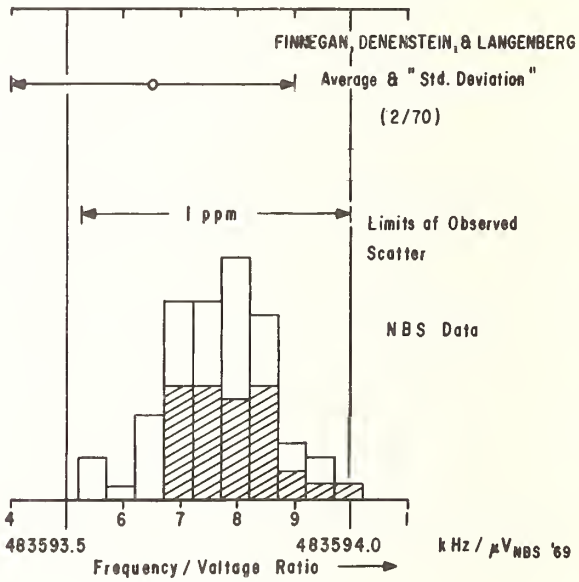


FIGURE 4. *Histogram of the data in figures 2 and 3, sorted into 10 bins each 0.10 ppm wide.*

Ordinate corresponds to frequency of occurrence. Open blocks correspond to points of figure 2; hatched blocks, to those of figure 3.

Figure 4 is a histogram of these determinations, which is intended primarily to show a measure of experimental scatter. A one-ppm band, and the recent determination by Finnegan et al. [3] are indicated in the graph.

No average should be drawn directly from this histogram, since we feel that the points of different runs should carry differing weights. All of the points in this histogram were measured with like conditions

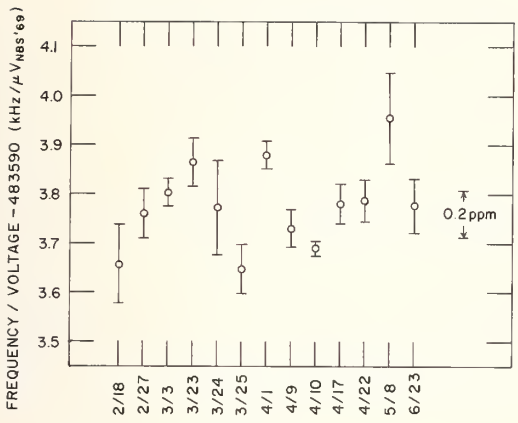


FIGURE 5. *Chronological plot of Means and Standard Deviations for daily runs shown in figures 2 and 3.*

Standard deviations are from observed scatter only, and do not include systematic uncertainties known to be present in the potentiometer. Variable errors in the potentiometric procedure, as well as errors in the knowledge of working-cell voltages (which include scatter in the assignments against the National Reference group), contribute to the observed scatter.

for one particular Kelvin-Varley divider; and we believe that it does provide some upper bounds for experimental random scatter in this combination of divider-and-Hamon-pair.

In figure 5 we show a chronological plot of the mean, and standard deviation from each day's run, as evaluated from the observed values in figures 2 and 3. Grouping by day's run corresponds to grouping by like Josephson-junction conditions, and like working-cell temperatures. Overall agreement of the runs with each other and with a common mean appears acceptable, if we add to each day's observed scatter a root-sum-square total of known experimental uncertainties.

The largest of these uncertainties cluster around the use of the Kelvin-Varley divider; and this is about what we might expect, since these devices are not specifically engineered for current division to parts in 10^7, at voltage levels between 2 and 10 mV. Power and temperature effects in the calibration; contact resistances; aging; and random scatter all contribute. Insertion of the voltage divider brings in a lead correction (see l_1 and l_2, fig. 1). In the present results, these are the price which has been paid for optimum signal-to-noise stability in the Josephson device itself.

To set some experimental limits on these uncertainties, we have tried substituting a Kelvin-Varley divider of different design and construction. Comparable within-run (short-term) stability, calibration accuracy, and current division have been observed with this divider. But we might expect to find some difference in the average observed value of $2e/h$, which we could use as a measure of systematic errors, known to be present. However, the observations taken with this second divider have so far been of poorer quality (for reasons quite unconnected with the divider) and have not yet permitted an accurate check of the difference.

We anticipate that a still better check of the series-to-parallel network may be made with the Kelvin-Varley divider removed, and the Josephson voltage fixed at 1/100 of the standard-cell voltage.

Statement of a final average or estimated systematic error, for the measurements presented in our graphs, is therefore reserved until these tests have been concluded in a satisfactory manner.

References

[1] Page, C. H., J. Res. Nat. Bur Stand. (U.S.), **69C** (Eng. & Instr.), No. 3, 181–189 (1965).
[2] Harris, F. K., Fowler, H. A., and Olsen, P. T., Metrologia, **6**, 134 (1970).
[3] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).

DISCUSSION

B. N. TAYLOR: I think rather than have any detailed discussion on this paper now we should wait until the other similar papers are given and then lump all the discussion into one. So if you have any questions, why don't you just jot them down.

Measurement of $2e/h$ by the Josephson Effect

B. W. Petley and J. C. Gallop

Division of Quantum Metrology, National Physical Laboratory, Teddington, Middlesex, England

The paper discusses recent work on the measurement of $2e/h$. Work has continued on the use of the solder-drop type of junction and the d-c measuring techniques have been improved, partly by using Hamon resistors.

Key words: Atomic constant; $2e/h$; electrical standards; Josephson effect; precise measurement.

1. Introduction

The work at the University of Pennsylvania [1, 2, 3] on the Josephson effect [4] has shown that it may be used to measure $2e/h$ with high precision and also that it may well provide a reliable quantum method of maintaining the volt [5]. The objective of our initial work [6] on the Josephson effect was to verify that there was not any major systematic error in the earlier work. A second objective was to ascertain the measurement problems posed in attaining the higher precision that is required in using the Josephson effects as a voltage standard. The agreement between our result and that at the University of Pennsylvania [2] was remarkable. The apparent closeness of the agreement highlighted two major problems. The first is the difficulty of estimating the systematic errors which dominate the measurement and the second is the difficulty of intercomparing the values when they are obtained in different countries where a different National volt is involved. There is at present no easy way of completely overcoming either of these problems.

At this stage of the work the precision is limited by the measurement process rather than by the Josephson effects themselves. We have continued to use for the present the solder-drop type of Josephson junction while we are developing the measurement techniques. Our cryogenic problems are then trivial since a normal laboratory helium storage dewar can be used as the cryostat [7] and one filling of 17 litres lasts for nearly two weeks. The solder-drop Josephson junction was devised by Clarke [8] and comprises a dry-soldered joint, of about an ohm room temperature resistance, between a short length of Niobium wire and a drop of solder through which it passes. Copper current and potential leads are immersed in the solder drop.

The current-voltage characteristic of a solder-drop junction when 36.4 GHz microwaves are incident on it is shown in figure 1. The current steps are separated by a potential difference that is $h/2e$ times the microwave frequency. The effects are reversible with current and the inevitable thermal emf may be eliminated by a simple reversal technique. Since the effect is independent of the step number the potential difference between a number of steps may be measured.

There remains a disparity between the potential to be measured and the potential against which it is to be compared of two or more orders of magnitude. The comparison of these disparate potentials is one of the factors limiting the precision of the measurement. This paper discusses some of the improvements that have been effected since our initial measurement in 1969.

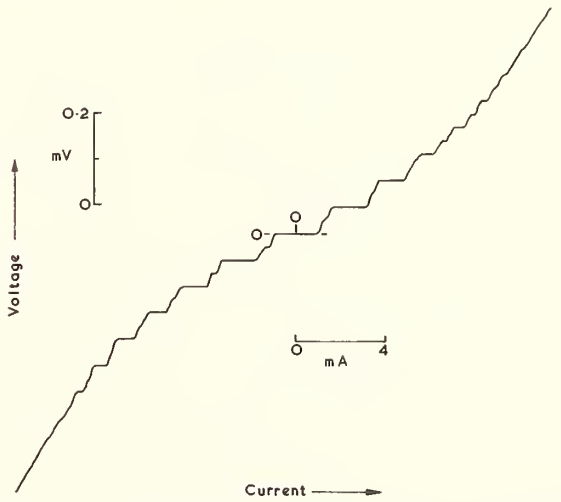


FIGURE 1. *The current-voltage characteristic of a solder-drop Josephson junction with 36.4 GHz microwaves incident on the junction* ($T = 4.2$K).

2. The Measurement Technique

The measurement of the ratio of the Josephson voltage to the emf of a Weston-cadmium cell can be

greatly simplified if the same current may be passed through two resistors connected in series. In such a case the cell emf may be balanced across one resistor and the Josephson voltage across the other, perhaps by changing the current through the resistors for the two balances by less than a factor of ten so that the current ratio can be measured precisely. The major ratio measurement is then that of a resistance ratio. A large resistance ratio is measurable to a much greater precision than a large current or voltage ratio.

There are several ways of measuring this resistance ratio to a part in 10^7. One is to use the Hamon principle [9, 2] for switching resistors from being connected in series to being connected in parallel, so that the resistance ratio is simply n^2 exactly. Another method uses the a-c ratio measuring techniques developed by Hill [10] and Deacon [11]. The first method suffers from the disadvantage that in Josephson effect work the resistors are not used at constant power, the power dissipated in the resistor changing a hundred fold in switching the resistors

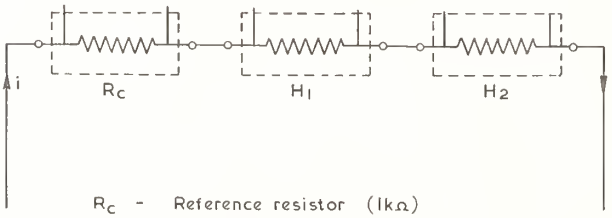


FIGURE 2. *The double Hamon resistor used to establish an approximate 900:1 resistance ratio to parts in 10^7.*

from being connected in series to parallel. There is also the further uncertainty that the thermal emf's in the complex network of resistors might not reverse exactly to the nanovolt level of precision when the current through the network is reversed. The second method has the inherent uncertainty that the a-c value might in some way differ from the d-c value. It appeared desirable therefore that both methods be investigated in one laboratory. We used the a-c method for measuring the resistance ratio in our initial work [6] and have now used a Hamon resistor.

The Hamon resistor is not very economical in its use of resistors to establish large ratios—thus a twenty element Hamon resistor would only yield a 400:1 ratio, whereas two ten-element resistors may be used to establish a 10^4:1 ratio with only a slight loss of precision. We have used a double Hamon resistor for the present series of measurements in order to establish a 900:1 ratio.

The measurement process is illustrated in figure 2. The standard cell is balanced across the resistor R_c and the Josephson voltage across the second Hamon resistor H2 with H2 connected in the parallel mode. The resistance ratio is then established by comparing the three element Hamon H1 in the parallel mode, with the ten element Hamon H2 in the series mode. The ratio of R_c to H1 connected in series is then established by a separate measurement.

Each element of the three-element Hamon, H1, consists of three 100 Ω resistors in series. The construction was thereby simplified, since readily available resistors could be used. The effects of any power coefficient of the resistance value were also reduced. In this way the ratio of the voltages could be established to a few parts in 10^7.

The screening of the room temperature part of the apparatus has been altered to suit the new equipment and the galvanometer amplifier that is used to balance the Josephson voltage has been shielded with mu-metal. The galvanometer sensitivity has also been improved from 3 mm per nV to 11 mm per nV, by reducing the input resistance.

In our previous determination the divider resistors were used in association with a Stabaumatic potentiometer. This meant that a separate calibration of the potentiometer was required. We have now substituted an inverted Kelvin-Varley divider [12] for the Stabaumatic potentiometer, purely as a method of varying the current through the Hamon resistors. The actual potential difference across R_c in figure 2 was measured with a current comparator potentiometer [13]. The calibration of this potentiometer is known to parts in 10^7. The standard cell emf was also measured using the same terminals of the potentiometer.

Two reference sources of emf were used, a local temperature-controlled standard cell in our laboratory and a remote bank of 12 cells located in the room in which the NPL volt is maintained. These cells could be compared directly with the NPL volt. Two twin-screened cables were installed between the laboratories, which were 50 m apart. The measurement loop could therefore be closed from either laboratory in order to establish the volt transfer-errors, e.g., by measuring the cell emf with the potentiometer connected directly to the cell or to the cell through the cables.

In our previous measurement the microwave frequency was stabilised to the fourth harmonic of a 9.2 GHz solid-state frequency source. In the present work we have again used a 36 GHz klystron but the frequency has been locked using a CML phase-lock system which permits operation over a wide range of frequencies. The electrically quiet environment of our laboratory in Bushy House at the NPL has enabled us to continue to operate with well screened apparatus without the need to use a screened room.

All equipment within 3 m of the junction is dc operated from stabilised storage batteries. These batteries are charged nightly to be at the beginning of their discharge plateau when work commences. The galvanometer lamps are operated from well stabilised mains-powered dc supplies.

The current-comparator potentiometer is not connected to the resistors while the junction potential

is being balanced and mu-metal screened chokes have been inserted in the potentiometer leads. The insulation resistance to ground of all equipment is checked to ensure that it is greater than 10^{10} Ω.

3. Results

The equipment has been operational for only a few months and only a few junctions have yet been measured. Previous work at the University of Pennsylvania and to a lesser extent by ourselves, has established that the value of $2e/h$ does not depend on the junction type or material at our present level of precision. However we have made measurements on three junctions and these have agreed to within 0.2 ppm. The standard deviation of a single set of measurements is now 1 nV.

Our experiments have shown that solder-drop junctions do not show steps far beyond 1.5 mV when operated at 4.2 K. The maximum step voltage does not depend markedly on frequency over the range 18 to 72 GHz.

Although an increase in the maximum observable step voltage is clearly desirable if the measurement precision is to be improved using present techniques, the solder-drop junction possesses the advantage of ease of production and is reasonably robust. One of our junctions has been transported to London twice between $2e/h$ measurements and the storage lifetime at 77 K is at least six months.

The overall precision of our measurements has therefore been improved by a factor of three during the last year and the standard deviation now corresponds to 0.8 ppm. Our provisional result in terms of the NPL as-maintained volt is that $2e/h$ is 483.5942± 0.00039 (0.8 ppm) MHz/μV NPL 69.

Our present result is close to our 1969 value being 0.2 ppm higher.[1] The result of the 1970 BIPM comparison of the volt have not been received at the time of writing and so we have not expressed our results in terms of BIPM or NBS volts.

[1] Measurements since the conference revealed a further correction which has now been included in the above number.

4. Acknowledgment

This work forms part of the research programme of the National Physical Laboratory.

5. References

[1] Parker, W. H., Taylor, B. N., and Langenberg, D. N., Phys. Rev. Letters **18**, 287 (1967).
Langenberg, D. N., Parker, W. H., and Taylor, B. N., Proc. Third Int. Conf. on Atomic Masses, R. C. Barber, Ed. (Univ. of Manitoba Press), 439 (1968).
[2] Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177**, 639 (1969).
Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).
Denenstein, A., Finnegan, T. F., Langenberg, D. N., Parker, W. H., and Taylor, B. N., Phys. Rev. **B1**, 4500 (1970).
[3] Finnegan, T. F., Denenstein, A., Langenberg, D. N., McMenamin, J. C., Novoseller, D. E., and Cheng, L., Phys. Rev. Lett. **23**, 229 (1969).
[4] Josephson, B. D.:
(i) Phys. Lett. **1**, 251 (1962);
(ii) Adv. Phys. **14**, 419 (1965).
[5] Taylor, B. N., Parker, W. H., Langenberg, D. N., and Denenstein, A., Metrologia **3**, 89 (1967).
[6] Petley, B. W., and Morris, K., Phys. Lett. **29A**, 289 (1969).
Petley, B. W., and Morris, K., Metrologia **6**, 46 (1970).
[7] Petley, B. W., and Morris, K., J. Sci. Instrum. (J. Phys. E), Series 2, **2**, 649 (1969).
[8] Clarke, J., Phil. Mag. **13**, 115 (1966).
[9] Hamon, B. V., J. Sci. Instrum. **31**, 450 (1954).
[10] Hill, J. J., Proc. IEE **112**, 211 (1965).
[11] Deacon, T. A., Proc. IEE **117**, 634 (1970).
[12] Julie, L., IEEE Trans. Instrum. Meas. **IM-16**, 187 (1967).
Abramowitz, Rev. Sci. Instrum. **38**, 898 (1967).
[13] MacMartin, P. M., and Kusters, N. L., Trans. IEEE Instrum. Meas. **IM 17**, 263 (1968).

Progress Toward the Josephson Voltage Standard: a Sub-Part-Per-Million Determination of $2e/h$ *

T. F. Finnegan, A. Denenstein, and D. N. Langenberg

Department of Physics and Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, Pa. 19104

An ac Josephson effect determination of $2e/h$ is described, in which an adjustable (via frequency) Josephson source voltage near 10 mV is compared with a standard cell voltage by means of a fixed 100:1 voltage ratio. Two methods for establishing this ratio have been used. One is based on a double series-parallel exchange resistance network, and the other on a double binary resistance chain which is self-calibrated. The ratios obtained with the two methods agree within 0.06 ppm. The precision of the measurements referred to the local voltage standard is better than 0.04 ppm. The accuracy of the measurement referred to the National Bureau of Standards (NBS) as-maintained volt is 0.16 ppm and is presently limited by uncertainties in relating the local volt to the NBS as-maintained volt. The demonstrated intrinsic precision of better than 0.04 ppm makes practical the ac Josephson effect as a fundamental solid state voltage standard. The techniques developed in the present measurements also provide sufficient stability and resolution to permit unambiguous studies of the factors which presently limit the reliability of conventional electrochemical voltage standards.

Key words: AC Josephson effect; standard cell; voltage comparator; voltage standard.

Two essential criteria which a primary voltage standard must satisfy are: (1) A phenomenon must exist (such as the ac Josephson effect) which develops a voltage sufficiently predictable and independent of all external parameters. (2) Technology must exist which permits comparison of this voltage with secondary standards (such as a conventional standard cell) at the required level of accuracy. The question of whether the ac Josephson effect satisfies the first of these two criteria has been investigated in considerable detail, both theoretically and experimentally. Neither theory nor experiment has yet uncovered or even suggested a correction at the level of 0.01 part-per-million (ppm) or $1/10^8$. The main purpose of this paper is to demonstrate that the second of these two criteria, i.e., establishment of the technology required to compare a Josephson device with a standard cell to better than a tenth ppm ($1/10^7$), has indeed been satisfied. We have recently reported preliminary results of a determination of $2e/h$ with a total uncertainty of about one half ppm [1]. Most of this uncertainty was associated with the voltage standard transfer between the University of Pennsylvania and the National Bureau of Standards. Because the voltage standard is the principal factor limiting the accuracy of our experiments, we discuss our standard cell maintenance and transfer in some detail in this paper.

In our earlier paper [1] we reported some of the advantages of increasing the Josephson device voltage from 1 mV to 10 mV; they include, in particular, a factor of ten increase in precision. Both a single junction and three connected in series were used in those measurements. Figure 1(a) shows the I-V characteristic of a lead-lead oxide-lead Josephson tunnel junction. Figures 1(b) and 1(c) show the characteristic on a reduced scale, first without and then with an applied microwave field. Figure 1(d) shows the microwave induced (Shapiro) steps on an expanded scale near 10 mV. The step amplitudes are approximately 20 μA. Since the maximum amplitude of these steps decreases as the bias voltage is increased, the use of a multiple junction device to achieve 10 mV allows the biasing of each junction on a relatively large amplitude step (50 to 200 μA), minimizing the uncertainty due to possible non-vertical steps caused by noise. Individual junctions in a multiple junction device can be compared in a differential experiment to demonstrate that the Josephson frequency voltage ratio does not depend upon the voltage level (step number). We have carried out such an experiment which indicates independence of the voltage level (between 2.5 and 5.0 mV) to better than 0.02 ppm ($2/10^8$). (In addition, the effects of thermal emfs in the leads from the cryostat may be investigated under operating conditions in the presence of the applied microwave field by suitably biasing the junctions in series opposition.) Such Josephson devices may be recycled and used essentially indefinitely when stored at liquid nitrogen temperatures. A single device has been kept in continuous use for some seven months during which time it was cycled to liquid helium tempera-

* Supported by the National Science Foundation and The Advanced Research Projects Agency.

tures some twenty-five times without any apparent change in the I-V characteristic.

The basic circuit diagram of the type of voltage comparison apparatus used in these measurements is shown in figure 2. The general operation of such a comparator, which makes use of a *fixed* voltage (resistance) ratio, consists of two steps. First, the working current of the comparator is adjusted so that the voltage across the 1 kΩ resistor is equal to that of the standard cell using the corresponding null detector (ND). Next, the frequency of the microwaves is adjusted to make the total junction

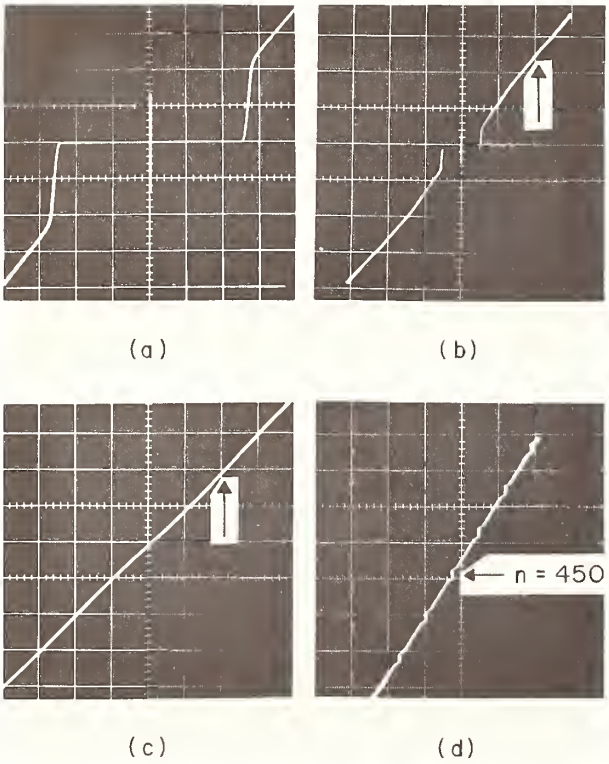


FIGURE 1. I-V *characteristic of a* Pb *Josephson tunnel junction.* (a) Vertical scale is approximately 2.5 mA/cm and horizontal scale is 1 mV/cm; (b) 10 mA/cm and 5 mV/cm; (c) also 10 mA/cm and 5 mV/cm; (d) 50 μA/cm and 25 μV/cm. The arrows in (b) and (c) indicate the voltage at 10 mV; the arrow in (d) indicates the induced step at about 10.2 mV corresponding to n = 450. The applied microwave frequency in (c) and (d) is approximately 11 GHz.

voltage equal to that developed across the 10 Ω resistor using the other ND. Now the relation between the standard cell voltage V_s and the microwave frequency, ν, is

$$\beta N\nu = (2e/h)\, V_s \qquad (1)$$

where β is the comparator ratio (in this case, 100) and N is the total device step number. It is convenient to define the left side of eq (1) as F_s, the "standard cell Josephson frequency."

A simplified circuit diagram of the double series-parallel exchange comparator we have built [2] is shown in figure 3. The 100:1 ratio is established using two series-parallel networks [3] having ten resistors each. The working current is supplied by two mercury batteries connected in series and regulated by a third mercury battery shunting the midpoint of the entire network. The dropping resistors are adjusted for negligible current in the regulating battery to attain the high stability of an unloaded mercury battery (about 0.5 ppm/h).

Since it was necessary to use low valued resistors in the series-parallel networks, some compensation is required to achieve high ratio accuracy with conventional switches. This compensation (not shown in fig. 3) took the form of tetrahedral junctions [4, 5] between the 100 Ω main resistors and compensating resistors (fans) in *both* the current and voltage leads [6]. The push-button indicated in figure 3 was used to introduce a known increment

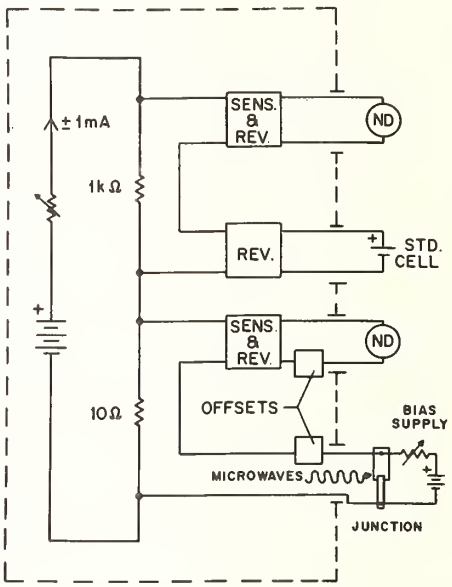


FIGURE 2. *Basic circuit diagram of the type of voltage comparison apparatus used in 2e/h measurements.*

of about 2 ppm in the working current for calibration purposes.

In practice, it was found convenient to balance the two null detectors to only about 1 ppm and to interpolate the remainder. To facilitate this interpolation and provide a permanent record, all ND balances were continuously recorded. Figure 4 shows a section of one of these recordings. The standard cell balance is shown in the upper part of the figure, and junction balance in the lower part. Because the power supply variations took the form of a uniform drift, it was adequate to make the two balances consecutively using a single ND and recorder rather than simultaneously with two of each. This required the application of a small drift correction. For each balance, the "+" and "−" refer to the polarity of the ND switch. The pulse at the center labeled "MS" is the calibrating signal (produced with the push-button

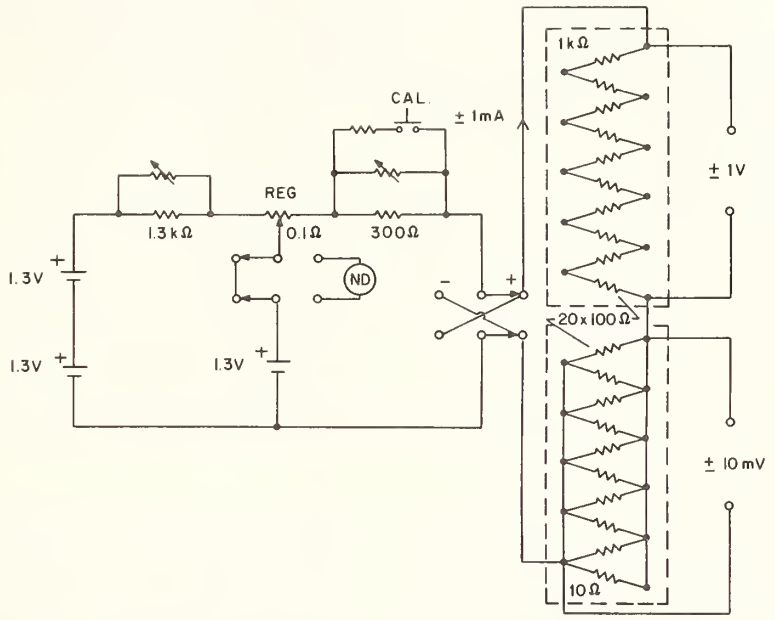


FIGURE 3. *Simplified circuit diagram of the double series-parallel exchange comparator.*

in figure 3) which enables accurate interpolation. The resolution of the standard cell balance was about 0.01 ppm ($1/10^8$); and that of the junction balance, about 0.02 ppm ($2/10^8$). In order to obtain a data point, it was necessary to make four pairs of balances of the type shown in figure 4. Two of these involve a reversal of all voltages to eliminate the effects of thermal emfs. Two involve exchanging the two series-parallel networks in order to eliminate the effect of inequality between the total resistance in each network.

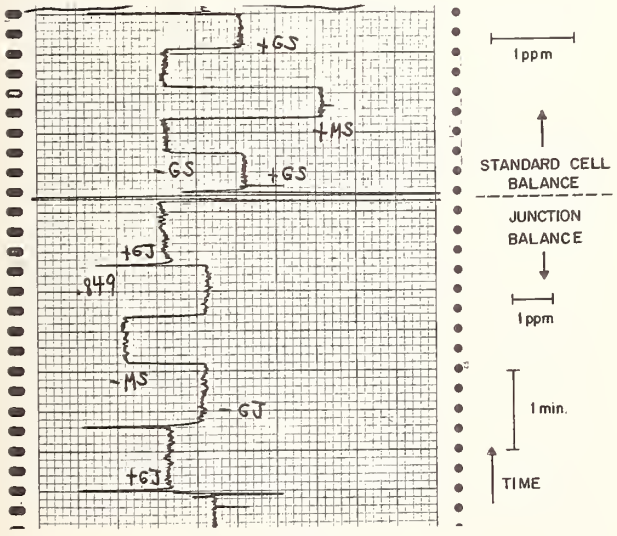


FIGURE 4. *Section of null detector recording showing standard cell and junction balances.*

The data from some recent runs are shown in figure 5. The quantity plotted is the standard cell Josephson frequency, F_s, versus time within a run. The same standard cell was used for each of these measurements. F_s is plotted because this is the experimentally determined quantity. The three runs shown were made at approximately one week intervals. The standard deviation of the individual points within a run ranged from 0.01 ppm to 0.05 ppm (1 to $5/10^8$); the standard deviations of the means were about half these values. The progressive drift of the data from week to week can be interpreted as drift in that cell, which had recently been transported.

Immediately before and after every run, the cell used was compared with the other members of its group. Figure 6 shows the regular intercomparison data of the two particular cells (each about the mean of its group) used in the $2e/h$ measurements reported here. Data over a hundred day period are shown. The random error of our standard cell comparisons was about 7 nV ($7/10^9$) for a *single* observation. All of our standard cell comparisons were made with cells in air baths. Considering the data for cell B3 (fig. 6(a)), it is apparent that most of the time the day to day fluctuations were about 10 nV ($1/10^8$). Occasionally, however, there was an abrupt change, indicating the unpredictability typical of some of our standard cells. This group of cells was not transported or disturbed during the time period indicated. Figure 6(b) shows the data for a second group which was transported to and from the National Bureau of Standards (NBS) for calibration during the time period shown.

Comparisons were also made between these two groups (A and B). Figure 7(a) shows the voltage difference between the group means ($\bar{A}-\bar{B}$) versus

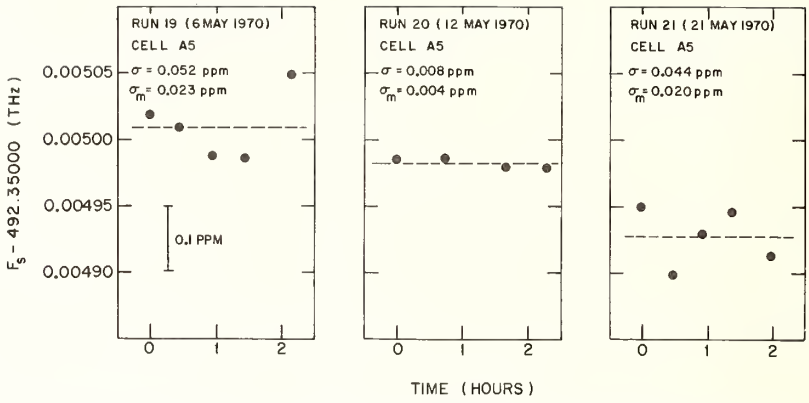


FIGURE 5. *Standard cell Josephson frequency, F_s, versus time for three runs.*
The standard deviations of the individual data points and of their mean are indicated for each run.

time. It is clear that the earlier data have considerably more scatter than the later data. Figure 7(b) and 7(c) show the standard cell Josephson frequencies ($F_s = (2e/h)V_s$) of the group means ($\bar{B}$ and $\bar{A}$ respectively) versus time. Each of these figures can be interpreted as a plot of $2e/h$ versus time, using the indicated group of standard cells as the reference voltage. Data from eleven runs are presented. For each run, the group containing the actual cell used is indicated by a solid point, i.e., group B for the runs numbered 11 through 18, and group A for runs 19 through 21. The open points were obtained from the standard cell intercomparison data in order to monitor the long term stability of both groups. Note that the earlier points in figures 7(b) and 7(c) do not show the scatter which might be expected from the data in the corresponding time period in figure 7(a). The solid points in figure 7(a) correspond to the individual runs and indeed do not show the large scatter characteristic of the remaining points in the same period. The reason for this appears to be that most of the runs during this time were made on weekends when (fortunately) the ambient conditions were most stable. The error bars on the individual points in figures 7(b) and 7(c) include the random error in the mean of the run, typically 0.02 ppm ($2/10^8$), plus an estimate of the uncertainty caused by the disturbing effects of the run itself on the ambient temperature, which particularly affected the standard cells. The standard cells contributed nearly all the latter uncertainty, which ranged from 0.02 to 0.05 ppm (2 to $5/10^8$). Both the standard cell comparison data and the "$2e/h$" data indicate that the mean of box A before transport to NBS and the mean after its return did not differ by more than about 0.2 μV.

Concurrently with some of the later measurements in the series, the University of Pennsylvania participated in the NBS Volt Transfer Program [7] in order to improve the tie between the respective voltage standards. This program requires the participating laboratory to "calibrate" a group of standard

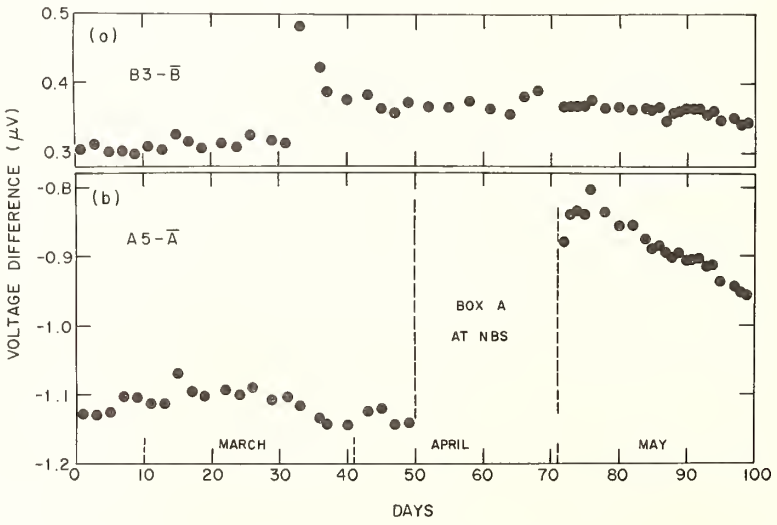


FIGURE 6. *Intercomparison data for the two particular cells (each about the mean of its group) used in 2e/h measurements.*

cells supplied by NBS. From these data, a voltage assignment is made by NBS to the laboratory's voltage standard. Figure 8 shows our data; group D was the one supplied by NBS. (The dashed lines are least squares fits.) In figure 8(a) the difference between the means of the two groups is plotted versus time. The standard deviation of these data is 0.043 μV; the data show no significant drift. Some of the scatter is due to the nonnegligible sensitivity of the temperature sensors in both groups to the ambient temperature. This effect is illustrated in figure 8(b), where the same data are plotted versus ambient temperature. The dependence on ambient temperature is clearly significant. Taking this dependence into account reduces the standard deviation to

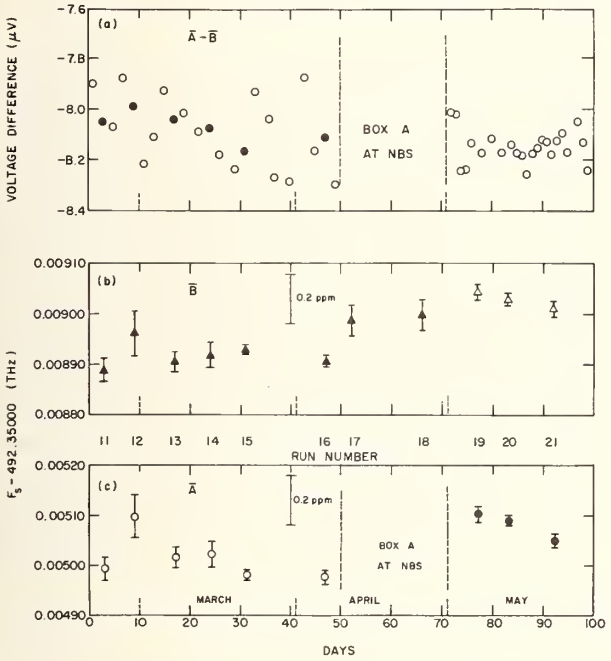


FIGURE 7. (a) *Voltage difference between the group means* $(\bar{A}-\bar{B})$ *versus time,* (b) *standard cell Josephson frequency,* F_s, *of the group mean* $\bar{B}$ *versus time,* (c) F_s *of* $\bar{A}$ *versus time.* All three parts have been plotted on a common time scale.

0.030 μV. This indicates that temperature corrections applied in the usual manner are not adequate.

Table 1 lists the principal sources of uncertainty for the data taken during the 100 day period considered here. The first part includes the contributions exclusive of the voltage standard. Item (a) is a weighted mean of a series of runs like those shown in figure 5. Items (c) and (d) were periodically checked by "going inside" the comparator and making direct measurements on the series-parallel networks. Item (f) is important because the power in the main comparator resistors is changed by a factor of 100 in going from the series to parallel connections. These heating effects were estimated by extrapolating from tests made at powers greater than the maximum operating power. Item (h) is the uncertainty in the power supply drift correction mentioned earlier. Because extensive guarding was used throughout the measuring circuit, it was possible to measure directly the various leakage resistances which contribute to item (j). Combining these measurement uncertainties in quadrature results in an uncertainty of about 0.04 ppm ($4/10^8$). This is a one standard deviation estimate of how well a perfectly predictable voltage standard could be calibrated with this system. Item 2 is an estimate of how much the mean of our standard cells departs from this ideal during a typical run, and, when combined with the total uncertainty of item 1, is a measure of how well our standard cells can be maintained using the ac Josephson effect.

TABLE 1. *Principal Sources of Uncertainty* (1σ)

1. Measurement uncertainties	*Parts in* 10^8
(a) Random error of the mean.	1
(b) Frequency stability.	1
(c) Main resistor mismatch.	0.4
(d) Fan resistor mismatch.	1
(e) Transfer resistances of tetrahedral junctions.	0.4
(f) Main resistor heating effects.	3
(g) Comparator temperature stability	0.3
(h) Working current stability.	0.5
(i) Calibrating signal accuracy.	1
(j) Leakage resistances.	1
(k) Effects of thermal emfs.	0.5
(l) Possible step bias current dependence	0.4
RSS Subtotal	3.9
2. Short term local volt stability	5
RSS Subtotal	6.4
3. Volt transfer to NBS [VTP (1970)]	15
RSS Total	16

In order to obtain a value of $2e/h$ in terms of a widely distributed voltage unit such as the NBS volt, an additional uncertainty for the transfer (including the effects of physical transport, comparison measurements, and short term cell stability) of about 0.15 ppm ($15/10^8$) [8] must be added to the previous ones as indicated in item 3. This gives an overall uncertainty in $2e/h$ of about 0.16 ppm ($16/10^8$) for our measurements expressed in 1970 NBS units. Because our most recent data have not yet been fully analyzed, including the results of the indicated voltage transfers to NBS, we cannot report a final value of $2e/h$ at this time. However, a tentative value based on data presented here (particularly the three measurements in May which coincided with the NBS Volt Transfer Program standard cell comparisons) is about 0.1 ppm larger than our recently published value [1].

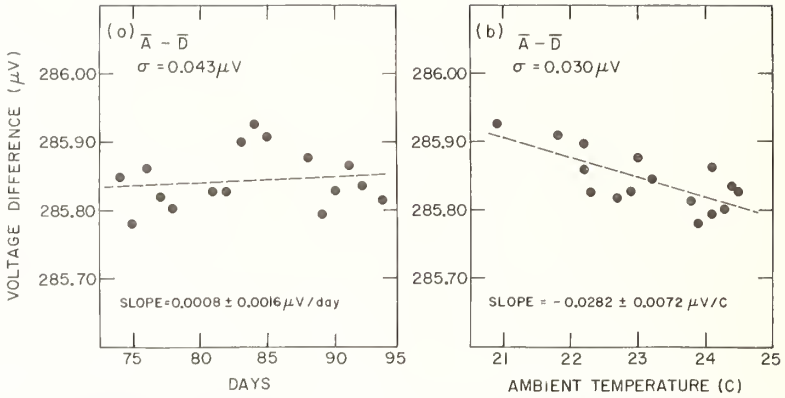


FIGURE 8. *Volt Transfer Program standard cell comparison data taken at the University of Pennsylvania.*
Group D was the one supplied by NBS. $\overline{A} - \overline{D}$ has been plotted versus time and ambient temperature.

Although the precision of our instrument appears to be quite good, there is still the possibility that a significant systematic error may have been overlooked. In order to test for such an error, a second instrument based on an entirely different principle for establishing the critical 100:1 voltage ratio was designed and built [9]. Figure 9(a) is a simplified circuit diagram of this second comparator in the measurement mode. The required voltages, 10 mV and 1 V, are developed using a voltbox in the same manner as in the series-parallel comparator (SPC). In operation, only two pairs of balances are necessary, in contrast to the SPC where four pairs of balances are made to include the network exchange. However, in the new method a calibration procedure is required [10]. A duplicate voltbox and power supply are included in the comparator for this purpose. Figure 9(b) shows the calibration mode in which the 100:1 ratio is established by high accuracy resistance comparisons. This is done as follows: Each resistor has a trimmer to allow fine adjustment of its value. With the switches set as shown (fig. 9(b)), a 10 Ω equal-arm Wheatstone bridge is formed. These resistors are trimmed until the ND indicates a balance for both positions of the "interchange" switch. The pair of 10 Ω resistors in each voltbox are then equal. The next position of the "select ratio" switch forms a 20 Ω bridge with the two 10 Ω resistors in each of the lower arms. Repeating the same procedure, the 20 Ω resistors are matched to the sum of the two 10 Ω resistors. Succeeding positions of the "select ratio" switch permit matching the remaining resistors to some previously measured combination until, on the last (seventh) calibration step, the 360 Ω resistor is matched to the sum of the 320 Ω and 40 Ω resistors. Now the resistors in each voltbox have the relative values indicated in figure 9. This technique is termed "cascade interchange." Lead compensation, required by the low resistor values employed, is provided by use of an independent power supply for each voltbox. It is important to note that both of our instruments are entirely self contained and require only an external null detector for all calibration and measurement operations.

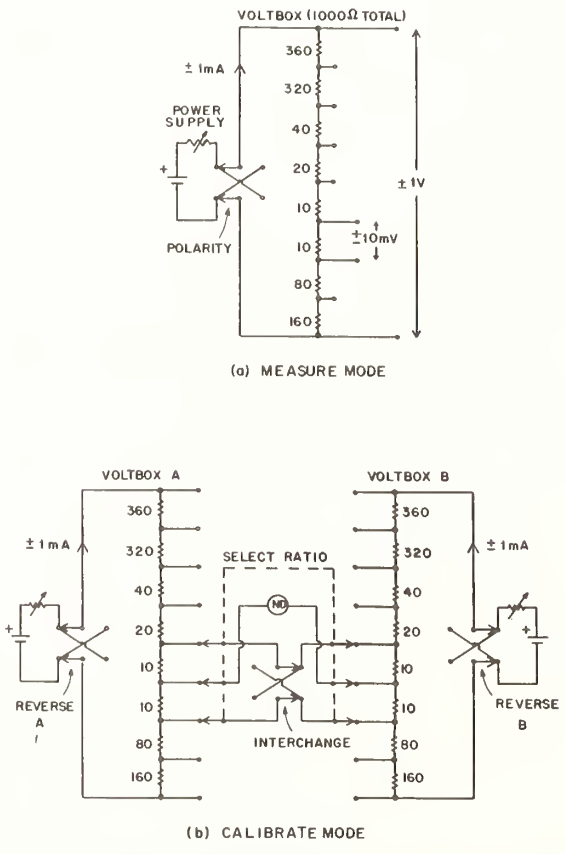


FIGURE 9. *Simplified circuit diagram of the cascaded interchange comparator.*
(a) measure mode; (b) calibrate mode. All resistor values are in ohms.

These two instruments have been compared in some recent runs by measuring $2e/h$ with each comparator during the same run. The *a priori* estimated accuracy of the cascaded interchange comparator (CIC) was comparable with that of the SPC. Preliminary analysis of the data from these runs shows agreement between the two comparators to better than the combined uncertainties, i.e., about 0.06 ppm

$(6/10^8)$. This close agreement is strong evidence that the estimated uncertainties peculiar to each instrument are realistic. For example, the leading uncertainty of the SPC (resistor heating) is not present in the CIC since in the latter the voltboxes are kept under the same power in both modes (see fig. 9). On the other hand, the first order calibration uncertainty in the CIC is not present in the SPC, which does not require calibration in the usual sense. In conventional methods, the critical ratios are calibrated relatively infrequently and not necessarily under the same conditions as in actual use. In contrast the high accuracy achievable with the series-parallel and cascaded interchange techniques results from establishing the critical ratio during actual use. Thus, any long term drift in our data should be assigned to the standard cells.

To summarize, then, we have demonstrated that it is possible to compare an rf-induced electrochemical potential difference across a Josephson device and the electrochemical potential of a standard cell with an uncertainty limited primarily by the stability and reproducibility of the standard cell. Apart from the obvious implications for the feasibility of a Josephson voltage standard, our measurements also have implications for a better knowledge of the fundamental constants.

When measured in the same electrical units $2e/h$ and γ_p, the gyromagnetic ratio of the proton, can be combined with other very accurately known (several tenths ppm or better) fundamental constants to determine a value of the fine structure constant α [11]. At present, γ_p (in NBS units) is known only to about 3 ppm. A measurement of $2e/h$ with an uncertainty of several tenths ppm is not only accurate enough to make the $2e/h$ contribution to the present uncertainty in α negligible, but is also sufficiently accurate so that any forseeable improvement in the measurement of γ_p (using the low field method) would yield an equivalent increase in the accuracy of α. Furthermore, if a Josephson device voltage were used to determine the current producing the magnetic field in the γ_p experiment, no standard cells with their associated uncertainties need enter into the experimental measurements determining α.

Note added in proof: The final result of the work discussed in this paper is $2e/h = 483.593\ 718 \pm 0.000\ 060$ $\mathrm{MHz}/\mu\mathrm{V}_{\mathrm{NBS69}}$ (see T. F. Finnegan, A. Denenstein, and D. N. Langenberg, Phys. Rev., in press).

Acknowledgments

We would like to thank the Electricity Division of the National Bureau of Standards, particularly W. G. Eicke, C. H. Page, and B. N. Taylor, for calibrating our standard cells and for the opportunity to participate in the Volt Transfer Program.

References

[1] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).
[2] Finnegan, T. F., and Denenstein, A., to be published.
[3] Wenner, F., J. Res. Nat. Bur. Stand. (U.S.) **25**, 229 (1940).
[4] Hamon, B. V., J. Sci. Instr. **31**, 450 (1954).
[5] Riley, J. C., IEEE Int. Conv. Rec., Part 11, 1965.
[6] Page, C. H., J. Res. Nat. Bur. Stand. (U.S.) **69C**, 181 (1965).
[7] Volt Transfer Program Instructions (Electricity Division, National Bureau of Standards).
[8] Eicke, W. G., and Taylor, B. N., private communication.
[9] Denenstein, A., and Finnegan, T. F., to be published.
[10] Denenstein, A., Dissertation, University of Pennsylvania, 1969.
[11] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

A Determination of $2e/h$ Based on the AC Josephson Effect

I. K. Harvey, J. C. Macfarlane, and R. B. Frenkel

A value of $2e/h$ has been determined by the ac Josephson effect using a niobium point contact junction operated at the 1 mV level.

Comparison with a standard cell is achieved using a divider having a ratio of 1,000:1. This divider is calibrated regularly using a Hamon-type bank of resistors. A servo control system sets the voltage across a 1 000 Ω resistor to equality with the standard cell voltage to high precision. The voltage from the junction is set to within 1 ppm of the voltage across a 1 Ω resistor by adjustment of frequency. Final balance is calculated from a continuous record of galvanometer deflection appropriately calibrated in terms of frequency. A record is also made of the servo control system error.

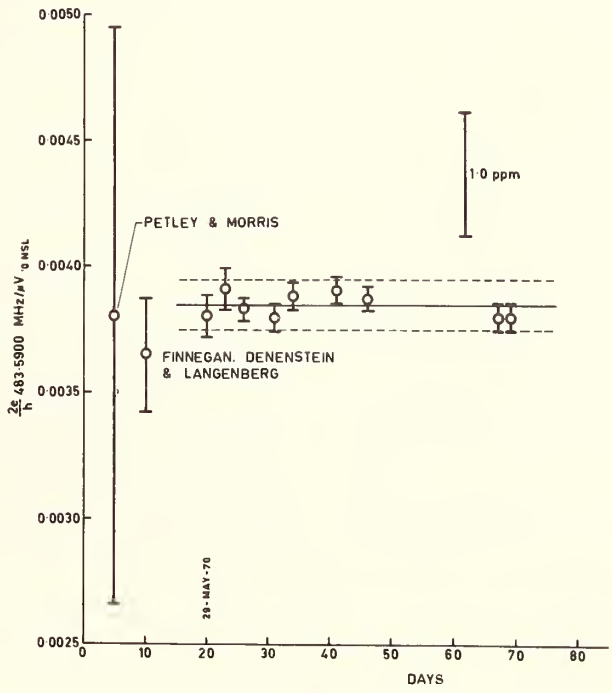
Thermal voltages are eliminated by a cancellation circuit and by reversal of divider current, standard cell and junction voltage.

Values of $2e/h$ were determined from measurements taken over a period of 50 days and are shown in the figure. The error bars were calculated as the root sum square of the standard error of the days' results together with all known and estimated errors both random and systematic. The standard error of the days' results ranged from 0.03 to 0.07 ppm.

The final value of $2e/h = 483.59384$ MHz/μV_{NSL} was calculated as the mean of the separate determinations. A recent international voltage intercomparison at BIPM indicated $V_{NSL} = V_{BIPM}$. An estimate of the standard deviation for this voltage intercomparison is indicated by the dotted lines.

Key words: Fundamental constants; Josephson effect; voltage standard.

Editors Note: For a report of this work, see Harvey, I. K., Macfarlane, J. C., and Frenkel, R. B., Phys. Rev. Letters **25**, 853 (1970).



Problems Concerning Josephson's Transient Effect for Comparisons of Standard Normal Elements

B. N. Oleinik, S. V. Gorbatzevich, K. A. Krasnov, and V. S. Umantzev

All Union Scientific Institute for Metrology, U.S.S.R.

Methods for comparisons of standard normal elements by using Josephson's transient effect are considered: a method using the divider to he regulated and a method using the generator to be retuned. The advantages and the disadvantages of the comparator circuits used with these methods are analyzed.

The requirements for separate units of the comparators—a microwave generator and a voltage divider—are given. As regards the method of the generator to he retuned, the limits of frequency instability, a range and smoothness of the retuning as well as the required power of the generator are determined.

An order of the division coefficient, the output resistance and the permissihle error of calihration of a voltage divider have been determined.

A traveling-wave tuhe generator stahilized hy a volumetric cavity is recommended as a generator which meets all the requirements mentioned ahove.

As regards the use of the comparators for the circuits under consideration, a printed-circuit resistor and a wire one have been compared, and the advantages of the former have heen shown. On account of the experience of the industrial manufacture of the printed-circuit resistors with an error of the order of 10^{-6}, the possibilities and perspectives of further improvement of accuracy of the printed-circuit resistors and the instruments on their hasis are shown. The data of stahility, temperature dependence of resistance and t.e.m.f. as well as the data of other parameters of the printed-circuit resistors are given. It is noted that the use of large integral resistive circuits permits to decrease an error of calihration of the printed voltage dividers to 10^{-8}. The prohlems of self-calihration of such dividers arc considered.

Inter-group comparisons of the normal elements present an urgent metrological prohlem for decision of which Josephson's effect can be used. It is noted that in case of using the industrial microwave generators and the voltage dividers there is a possihility of decreasing the error of comparisons hy two orders as against that ohtained at present time.

Key words: Josephson effect; printed circuit resistor; voltage divider.

DISCUSSION

B. N. Taylor: I will now open discussion on all of the experimental papers on the measurement of $2e/h$.

I would like to comment that I think that the accuracy of the $2e/h$ measurements has now reached the point where comparing values obtained in different laboratories by means of interlaboratory comparisons of standard cells is going to be futile. The transfer typically runs a few tenths of a ppm, and that is just about where we are now in $2e/h$. It seems to me that it is going to be quite difficult to say who is right and who is wrong, and probably the only way that the problem can be solved is to actually transport Josephson effect measuring apparatus from laboratory to laboratory and compare them against each other.

D. N. Langenberg: May I add one comment to the effect that if you really believe in the Josephson frequency-voltage relation at the level at which I think Doug Scalapino believes in it (and maybe I do), maybe what you ought to do is not transport the apparatus but to believe that the difference in experimental values of $2e/h$ really are beginning to reflect the uncertainty in the existing volt transfer process.

B. N. Taylor: That, of course, means that one has to have faith that everyone is equally competent in building Josephson apparatus.

Measurement of h/m_e Using Rotating Superconductors*

W. H. Parker and M. B. Simmonds**

Physics Department, University of California, Irvine, Calif. 92664

The fundamental constant h/m_e where m_e is the free electron mass has been determined from measurements on a rotating superconducting thin-film ring containing a single weak link. The value obtained agrees with the accepted value within the present uncertainty of 0.04 percent.

Key words: Compton wavelength of electron; electron mass; fundamental constants; Josephson effect; superconductivity.

1. Introduction

The possibility of measuring fundamental physical constants using the unique macroscopic quantum nature of the superconducting state has been strikingly demonstrated by the recent measurements of e/h using the ac Josephson effect [1]. These measurements have had a profound effect on many areas of physics, in particular the numerical values of the fundamental constants and the validity of quantum electrodynamics [2]. However, e/h is not the only fundamental quantity which can be measured using the properties of the superconducting state. This paper is a report on the progress of a series of experiments aimed at measuring the fundamental physical constant h/m_e where m_e is the rest mass of the free electron. This ratio can be obtained from measurements on rotating superconducting rings as described in section 2. The instrumentation and other experimental considerations necessary for successful measurements are discussed and preliminary data are presented in section 3. In section 4 we discuss the prospects for increased accuracy sufficient to contribute to the numerical values of the fundamental constants.

2. Quantum Coherence in Rotating

F. London in 1935 was the first to discuss the properties of superconductors in terms of a highly correlated, phase coherent quantum state [3]. This idea is of central importance in the phenomenological theory of superconductivity of Ginzberg and Landau and was given firm theoretical support by the microscopic theory of Bardeen, Cooper, and Schrieffer (BCS) in 1957 [4]. BCS showed that at sufficiently low temperatures, the electrons in certain metals are coupled together by a weak attractive interaction and form bound pairs (Cooper pairs). As a consequence of the considerable spatial overlap of these bound electron pairs, the phase of the wave functions of all these electron pairs must be identical in order to satisfy the exclusion principle and to minimize the total free energy of the metal. This is the microscopic origin of the macroscopic phase coherence which characterizes the superconducting state.

This quantum phase is related to the canonical momentum of the Cooper pairs by

$$\varphi(r_1, r_2) = \delta(r_2) - \delta(r_1) = \hbar^{-1} \int_1^2 \mathbf{p} \cdot d\mathbf{r} \qquad (1)$$

where φ is the phase difference between the two points $\mathbf{r}_2$ and $\mathbf{r}_1$ and $\mathbf{p}$ is the canonical momentum $m^*\mathbf{v} + e^*\mathbf{A}$. Here m^* and e^* are respectively the mass and the charge of the basic charged "particle" of the superconducting state. For a multiply connected bulk superconductor, the requirement that the wave function be single valued implies that the magnetic flux Φ cannot be arbitrary but must equal one of the values nh/e^* where n is an integer. The experimental discovery of the flux quantum in 1961 by Deaver and Fairbank [5] and by Doll and Näbauer [6] confirmed the existence of the phase coherent state. Numerically, e^* was found to be $2e$ as predicted by the BCS theory.

If the multiply connected superconductor is a simple-bulk ring and is rotated about its axis, then the quantization of the action requires

$$\begin{aligned} \hbar^{-1} \oint (m^*\mathbf{v} + e^*\mathbf{A}) \cdot d\mathbf{l} &= (m^*/\hbar) \oint r\omega dl + (e^*/\hbar)\Phi \\ &= (2m^*\omega S/\hbar) + (e^*/\hbar)\Phi \\ &= n2\pi \end{aligned}$$

or

$$\Phi/\Phi_0 = n - (2m^*\omega S/h) \qquad (2)$$

where $\Phi_0 = h/e^*$, S is the area enclosed within the ring, and ω is the rotation rate. Equation (2) is plotted in figure 1a where the magnetic flux as a function of the rotation rate is represented by a

* This work was supported by the National Science Foundation and the Alfred P. Sloan Foundation.

** Alfred P. Sloan Foundation Fellow.

family of curves, each member corresponding to a different value of n. These curves intersect the ω axis at values of ω differing by $h/2m^*S$. If by some means this separation of flux states along the ω axis can be measured along with the enclosed area S and rotation rate ω, then a value for h/m^* can be obtained.

For a bulk superconducting ring, the value of the integer n characterizing the trapped flux state is constant, thus prohibiting any measurement of the spacing between the states. In order for any possibility of measuring this separation to exist, transitions must be induced between the flux states. This can be accomplished by "weakening" the ring at a point such that the critical current of the ring becomes comparable to the current circulating in the ring.

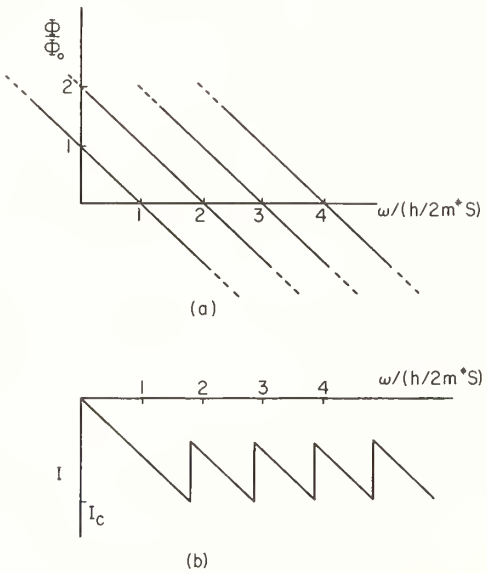


FIGURE 1. (a) *The possible states for the magnetic flux trapped in a bulk superconducting ring as a function of the rotation rate.* (b) *The behavior of the current in a superconducting ring containing a weak link as a function of rotation rate.*

A geometry particularly suitable for experiments on superconducting rings is shown in figure 2. A thin film of a convenient superconductor such as Sn is evaporated onto a fused silica rod to a thickness of a few hundred angstroms. The ring is "weakened" at one point by manually scribing a small constriction into the thin film approximately 1 μm wide [7].

Applying the quantization condition to this ring and remembering that the velocity of the electron pair consists of the rotational velocity $\mathbf{r}\times\omega$ plus the velocity due to current flow $\mathbf{j}/\rho e^*$ where $\mathbf{j}$ is the current density and ρ the electron pair density, one obtains

$$\oint[m^*\mathbf{j}\cdot d\mathbf{l}/\rho(e^*)^2] + \oint(m^*\mathbf{r}\times\omega\cdot d\mathbf{l}/e^*)+\Phi=nh/e^*=n\Phi_0. \quad (3)$$

The first integral in this equation is a function of temperature, thickness of the thin film, the dimensions of the weak link and even of the total current I. We can, however, define a function γ such that

$$\oint(m^*\mathbf{j}\cdot d\mathbf{l}/\rho(e^*)^2)=\gamma LI \quad (4)$$

where L is the inductance of the superconducting ring.

The second integral can be reduced using suitable vector identities to

$$\oint(m^*/e^*)\mathbf{r}\times\omega\cdot d\mathbf{l}=(2m^*/e^*)\,\omega\cdot\mathbf{S} \quad (5)$$

where S is the area enclosed by the ring regardless of the exact shape of the ring. Also, the enclosed flux is composed of an ambient flux Φ_A and a self generated flux LI. Equation (3) now becomes

$$\gamma LI+(2m^*\omega\cdot\mathbf{S}/e^*)+\Phi_A+LI=n\Phi_0$$

or

$$I=(\Phi_0/L_R)[n-(\Phi_A/\Phi_0)-(2m^*\omega\cdot\mathbf{S}/h)] \quad (6)$$

where $L_R=(1+\gamma)L$.

The current in the ring will increase negatively with increasing rotation rate ω until I equals the

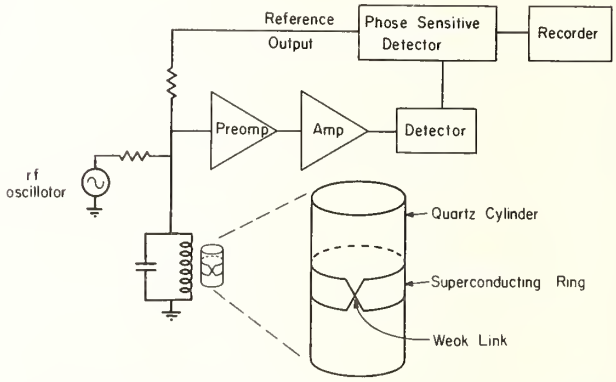


FIGURE 2. *Block diagram of the electronic system used to sense the current in a superconducting ring containing a weak link.*

critical current at which point the ring will undergo a transition ($\Delta n=1$) to an adjacent flux state [8]. As the rotation rate continues to increase, the current in the ring becomes periodic in the quantity $2m^*\omega\cdot\mathbf{S}/h$ as shown in figure 1b. If this periodic behavior of the current in a rotating superconducting ring can be measured, then the quantity h/m^* can be determined. A technique of observing this periodicity has been developed and is discussed in the next section.

However, before proceeding, it is essential to briefly discuss why the effective mass m^* entering eq (6) is expected to be exactly twice the mass of the free electron in the nonrelativistic limit. As in the measurement of e/h where the question was whether the complex many-body interactions in normal metals and superconductors might not lead to a

renormalization of the effective charge of the electrons, here the question is whether such interactions might not lead to a renormalization of the effective mass of the electrons. Although today there exists little doubt that the charge of the Cooper pair is twice the free electron value to an accuracy of at least 1 part in 10^7, the same cannot be said of the mass of the Cooper pair.

Before mentioning possible corrections to the mass, it must be pointed out that the concept of the "effective mass" used throughout solid state theory does *not* apply to the inertial properties of electrons in metals. It is a concept used to describe the dynamic properties of electrons in externally applied electric and magnetic fields when these fields are not corrected for local electric and magnetic effects. For reasons of convenience, the effects of the crystal fields are treated as changes in the dynamic properties of the electron rather than as corrections to the applied fields.

In this experiment we are concerned with the inertial mass of the electrons and both the microscopic theory and general arguments based on Galilean translation invariance indicate that the mass m^* is exactly twice the free electron mass, neglecting relativistic considerations [9]. The measured value of h/m_e supports this contention. It is found that to within the present accuracy of 0.04 percent, the value of m^* is twice the electron mass while the solid state effective mass of Sn is approximately 1.1 m_e.

As these measurements of h/m_e become more accurate, it will become necessary to understand any relativistic corrections to the mass of electrons in metals. To date no detailed calculations of such corrections have been made. One correction that does come to mind is the relativistic dependence of mass on velocity. For typical metals, the velocity of the electrons at the Fermi surface is approximately 10^8 cm/s which implies a relativistic mass *increase* of approximately 10 ppm. It has been argued, however, that this is not the relevant correction but rather the work function of a metal W implies a decrease of the mass by W/m_ec^2 [10]. Since a typical value of the work function is a few electron volts, the rest mass of the electron in the metal may actually be *decreased* by approximately 10 ppm. In light of the uncertainty of even the sign of the correction, the only conclusion to draw at the moment is that relativistic corrections probably are of order 10 ppm, and the need exists for a detailed calculation.

3. Experimental Measurements

In order to determine h/m_e, several quantities need to be measured; these are the periodicity of the current in the ring as a function of rotation rate, the rotation rate itself, and the enclosed area. The periodic response of the current in a rotating superconducting ring containing a weak link can be observed using the circuit of figure 2. The superconducting ring is placed within the coil of a radio-frequency resonant L–C circuit driven by a constant current source at the resonant frequency. The ring is coupled to the coil of the rf resonant circuit by mutual inductance. An analysis of this circuit shows that the ring can be considered as an additional parametric impedance in the resonant circuit. As a result, the ring's periodic response to either magnetic flux or rotation is reflected in the amplitude of the rf voltage across the resonant circuit.

The details of the interactions that result in the rf voltage reflecting the periodicity of the current in the ring have been studied using an electronic analog computer [11, 12]. These analog studies have led to the identification of the relevant parameters and their appropriate values to maximize the signal due to the quantum phenomena in the superconducting ring. To further improve the signal to noise ratio, a magnetic flux modulation is applied to the superconducting ring and the resulting oscillation in the rf amplitude is detected with a phase sensitive amplifier.

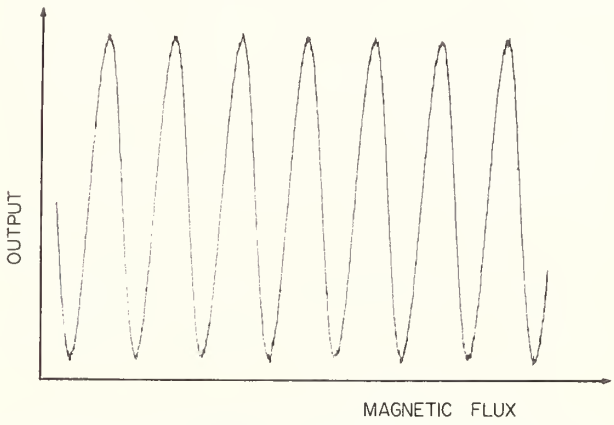


FIGURE 3. *Typical result for the output of the electronic instrumentation versus the magnetic flux through a 2 mm superconducting ring.*
An identical result is obtained if the output is plotted vs. the rotation rate. A 0.01 s time constant was used.

Typical experimental results are shown in figure 3 where the output of the electronic detecting system is plotted versus an applied dc magnetic field. As implied by eq (6), the result is identical when the detected signal is plotted versus the rotation rate. The amplitude and shape of this response curve are complicated functions of many parameters of the system but the periodicity is identical to that of the current in the ring.

The system is of sufficient sensitivity and stability that one oscillation of this response curve can be divided into 1000 parts. With the maximum rotation rate presently available (40 Hz) and rings approximately 3 mm in diameter, the number of oscillations observed is approximately 10, corresponding to a resolution of 100 ppm. However, this uncertainty is reduced to 50 ppm by measuring between $-\omega$ and

$+\omega$. If the diameter of the ring can be increased to 1 cm and the rotation rate to 100 Hz, the resolution in the periodicity measurement could be reduced to 2 ppm.

The fused silica cylinder is rotated by a synchronous motor driven by a power amplifier and a stable oscillator of known frequency. It has been verified that the synchronous motor maintains its phase relation with the drive signal to better than $\pm 2°$. Thus in one minute, the average frequency of rotation can be obtained to an accuracy of approximately 2 ppm.

Since the precision equipment and length standards needed for the area determination are not available in our laboratory, the National Bureau of Standards has kindly agreed to make the necessary measurements. Following standard metrological procedures, the diameter of the cylinders has been determined with a one standard deviation uncertainty of about 0.13 μm. However, a knowledge of a diameter (or any numbers of diameters) is insufficient to determine the area until the "out-of-round" condition of the cross section of the cylinder is determined. Measurements from the Talyrond instrument provide the information necessary to obtain the area from a single diameter measurement. With these measurements, the uncertainty in the area of a 2 mm cylinder at room temperature is about 0.015 percent.

A correction must be applied for the thermal expansion of fused silica from room temperature to 4 K. Since this total correction is at the moment smaller than other uncertainties, we have used $+80\pm 20$ ppm as a nominal value for the thermal expansion of the cylinders [13].

There are a number of other considerations that must be satisfied before an accurate determination of h/m_e is obtained. First, since it is the scaler product of the rotation vector and the vector normal to the enclosed area that enters eq (6), the axis of the quartz cylinder must be nearly parallel to the rotation axis. This angle has been measured and found to be less than 0.25° and the resulting error less than 10 ppm. To reduce this error to 1 ppm would require a measurement to $\pm 5'$; this is not difficult.

The ambient magnetic field must be sufficiently uniform in space and time to produce no observable effect. For rings 2 mm in diameter, this requires that the magnetic field not change in time by more than $6\times 10^{-13}T$ while for a ring 1 cm in diameter, the field must change by less than $3\times 10^{-14}T$. This stability can be obtained using a superconducting shield, although there is the difficulty of magnetic flux lines trapped in the shield occasionally moving, thus changing the magnetic field distribution within the shield.

The requirement that the magnetic flux through the ring not differ by more than approximately $\Phi_0/100$ for any angular orientation of ring dictates the spacial homogeneity needed in the magnetic field. To date, we have met this requirement for 2-mm rings by using a single cylindrical mu metal shield with coils at each end and a superconducting shield located symmetrically inside.

The best numerical value for h/m_e obtained from measurements on rotating superconductors and its one standard deviation uncertainty is

$$h/m = (7.2768 \pm 0.0028) \times 10^{-4} J \text{ s kg}^{-1}.$$

This result, obtained from a 2 mm diameter rod, is in agreement with the presently accepted value. The indicated uncertainty is the root sum square of 0.005 percent uncertainty in the rotation rate, 0.01 percent in the period change, and 0.04 percent in the area.

The large uncertainty in the area is the result of our inability to obtain suitable weak links on the actual rods measured by NBS before the time of this conference. We have had to relate the area of the rod on which rotating data were obtained to the area of those measured by NBS. All rods were cut from one initial long fused silica rod but insufficient knowledge of the taper of that rod has resulted in a factor of 3 increase in the uncertainty of the cross-sectional area.

4. Discussion

The accuracy to which h/m_e can be measured using rotating superconductors by the method described in this paper will ultimately be limited by the intrinsic thermal noise in the ring. The analog computer studies have indicated that the quantum signals will become obscured when the rms magnetic flux noise equals one-half the quantum of magnetic flux [11]. A simple argument based on the equipartition of energy suggests that the minimum flux noise is $(LKT)^{1/2}$ where L is the inductance of the ring. Numerically this means that $LKT \leq \Phi_0^2/4$ or $LT \leq 7.7\times 10^{-8}$ HK. Rings as large as 1 cm in diameter operating at a temperature of 4 K still satisfy this inequality.

The analog studies have also indicated that the critical current of the weak link should approximately equal Φ_0/L for best operation. Measurements of the current-voltage characteristics of weak links indicate that the minimum critical current not completely obscured by the intrinsic noise of the weak link is approximately 2×10^{-7} A [14]. This limits the inductance of the ring to $L \leq 1\times 10^{-8}$ H or to rings of diameter less than 7 mm. However, the detection system will still operate, with somewhat reduced signal to noise, for critical currents several times Φ_0/L. This criterion again indicates a maximum diameter of approximately 1 cm.

There are a number of difficulties that must be overcome before measurements can be made on 1-cm-diam rings. Requirements on the temporal and spatial homogeneity as well as the magnitude of the ambient magnetic field are increased by a factor of 10 beyond what we have presently obtained. Achieving such fields will be challenging but should not be impossible. Several promising techniques have been discussed in the literature [15].

In order to observe useful signals from 1-cm rings, the critical current of the weak link must be reduced by a factor of 10. In principle, this could be achieved by operating at temperatures sufficiently close to the transition temperature. However, for a number of experimental reasons, this does not appear possible. Several techniques which show promise of further "weakening" a weak link are being investigated. Two of these are (1) the use of an electron beam to reduce the dimensions of a bridge to several thousand Angstroms [16], and (2) the use of the proximity effect of a normal or ferromagnetic metal on a superconductor to weaken the coupling between two superconductors [17].

Assuming that solutions to these problems can be found, then using a 1 cm ring, both the periodicity of the current in the ring and rotation rate could be measured to an accuracy of about 2 ppm. Techniques under development at the National Bureau of Standards should provide measurements of the diameter of a high-quality fused silica cylinder to an accuracy of 15 nm or better [18]. Roundness measurements can be made to an accuracy of at least 5 nm [19]. Together, these measurements could provide an area determination accurate to 3–5 ppm at room temperature. Optical measurements of the thermal expansion of fused silica between room temperature and 4 K should contribute less than a few parts per million uncertainty to the area measurement.

Combining these experimental uncertainties, the present practical limit to the accuracy of a determination of h/m_e utilizing the periodic response of the current in a rotating superconducting ring containing a single weak link is 5–6 ppm.

The significance of such a measurement of h/m_e for the fundamental constants is best obtained from the relation

$$\alpha^2 = (2R_\infty/c)\,(h/m_e)$$

where α is the fine structure constant and R_∞ the Rydberg constant for infinite mass. A measurement of h/m_e to 6 ppm would imply a value for α accurate to 3 ppm, only twice the present uncertainty in α [2]. Thus at this level of accuracy, measurements on rotating superconductors would provide an additional input datum for a least squares adjustment of the fundamental constants.

5. Acknowledgment

We would like to thank Dr. J. A. Simpson of the National Bureau of Standards for arranging the measurements necessary to determine the area of the fused silica rods.

6. References

[1] Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177**, 639 (1969); Denenstein, A., Finnegan, T. F., Langenberg, D. N., Parker, W. H., and Taylor, B. N., Phys. Rev. B **1**, 4500 (1970); Petley, B. W., and Morris, K., Metrologia **6**, 46 (1970); and Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).

[2] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969). This paper has been reprinted as a monograph, The Fundamental Constants and Quantum Electrodynamics (Academic Press, New York, 1969).

[3] See for example London, F., Superfluids, Vol. 1 (John Wiley & Sons, New York, 1950).

[4] See for example DeGennes, P. G., Superconductivity of Metals and Alloys (W. A. Benjamin, Inc., New York, 1966).

[5] Deaver, B. S., Jr., and Fairbank, W. M., Phys. Rev. Letters **7**, 43 (1961).

[6] Doll, R., and Näubauer, M., Phys. Rev. Letters **7**, 51 (1961).

[7] Other techniques for obtaining a weak link include photoetching, electroetching, and electron beam etching.

[8] Experiments have shown that when the current in a superconducting ring equals the critical current, the flux changes by one quantum rather than by the number which would reduce the current to its smallest value consistent with flux quantization. See Silver and Zimmerman, Phys. Rev. **157**, 317 (1967). This reference also contains a discussion of the behavior of superconducting rings containing a weak link in applied magnetic fields.

[9] See the paper by D. J. Scalapino presented at this conference.

[10] See the comment from B. Josephson in the article by P. W. Anderson, in Progress in Low Temperature Physics, Vol. **5**, edited by C. J. Gorter (North-Holland Publ. Co., Amsterdam, 1967).

[11] Simmonds, M. B., and Parker, W. H., J. Appl. Phys. **42**, 38 (1971).

[12] These interactions have also been discussed by Zimmerman, J. E., Thiene, Paul, and Harding, J. T., J. Appl. Phys. **41**, 1572 (1970).

[13] Gibbons, D. F., J. Phys. Chem. Solids **11**, 246 (1959).

[14] Simmonds, M., and Parker, W. H., Phys. Rev. Letters **24**, 876 (1970).

[15] See for example Vant-Hull, L. L., and Mercereau, J. E., Rev. Sci. Instr. **34**, 1238 (1963); Deaver, B. S., Jr., Brown, R., and Hubbard, W., Rev. Sci. Instr. **36**, 1378 (1965).

[16] See for example Miyauchi, S., Tanaka, K., and Russ, J. C., Solid State Tech., p. 43, July 1969.

[17] Friebertshauser, P. E., Notarys, H. A., and Mercereau, J. E., Bull. Am. Phys. Soc. Ser. II, **13**, 1670 (1968).

[18] Simpson, J. A., private communication.

[19] Jurisson, J., Matthys, R. J., Plasek, F., and Thompson, H. J., Rev. Sci. Instr. **38**, 243 (1967).

DISCUSSION

J. A. Tyson: Have you tried shielding with superconductors?

W. H. Parker: Yes.

J. A. Tyson: What happens?

W. H. Parker: The trapped magnetic flux moves. In other words, if you observe the field within a superconducting shield at the level of 10^{-9} gauss, you see changes in the magnetic field. However, we think we understand some of the problems and we have some ideas of what to do about them. We will obviously continue working on this problem.

C. H. Page: I've got a suggestion that won't help you cut down the magnetic flux, but if you could put another one of these things in without rotating it,

this gives you a monitor as to whether something else happens.

W. H. PARKER: Yes.

E. R. COHEN: Is quartz the best cylinder to use? May there not be a better material?

W. H. PARKER: There are both advantages and disadvantages to quartz. I have been told that quartz is a very poor material to try to grind and polish because it has a tendency to chip and leave small holes in the surface of the rod. On the other hand, it probably has the minimum thermal expansion between room temperature and 4 kelvins. If we can find a way to both polish the rod and at the same time keep it round—I understand you can polish quartz, but to my knowledge it's not possible to both polish it and keep the rod round—then quartz is the best material. If that problem can be met by some manufacturer that we're unaware of, then I think quartz, or quartz with a slight addition of some impurity which may even reduce the thermal expansion, would be the material to use. If anyone knows of a manufacturer of quartz who can grind and polish it, I would appreciate knowing about it.

E. R. COHEN: Is it important to have small thermal expansion?

W. H. PARKER: It's convenient, in that we can tolerate a less precise measurement of the thermal expansion. In principle we can tolerate any magnitude of thermal expansion because it can be measured.

P. L. BENDER: It seems to me if you can measure dimensions at the 20 or so nanometer level at liquid helium temperature, then there would be a possibility of doing a high field proton gyromagnetic ratio experiment by taking a rectangular piece of fused silica and plating the coil around the edges of it and then being able to get the dimensions of that to a few parts in 10^7. So I think the problem of learning how to make accurate dimensional measurements at low temperatures is really important for several different measurements.

Influence on the Fundamental Constants of a Precise Measurement of the Josephson Effect in Liquid Helium

B. M. Khorana* and D. H. Douglass, Jr.

Department of Physics and Astronomy, University of Rochester, Rochester, N. Y. 14627

Using the analog of the a-c Josephson effect in superfluid helium, it should be possible to determine the product h/M_{He}, where h is Planck's constant and M_{He} is the mass of the helium atom, to an accuracy of a few ppm. The implications of such a measurement for the fundamental constants is discussed.

Key words: AC Josephson effect; fundamental constants; superfluid helium.

The ac Josephson effect in liquid helium manifests itself through the fundamental relation

$$nh\nu = \Delta\mu_0 = M_{He} g \Delta Z + \cdots = (M^*_{He}/N) g \Delta Z + \cdots,$$

where $\Delta\mu_0$ is the chemical potential difference between two baths of liquid helium separated by a weak link, ΔZ is the pressure head difference, M^*_{He} is the mass of helium atom in amu, ν is the frequency of an ac chemical potential radiation field, n is an integer and other constants have their usual meaning. Stationary behavior in the flow properties of superfluid helium have been observed when the condition $h\nu = M_{He} g(\Delta Z)_0$ is satisfied, and we have previously reported measurements in agreement with the expected value within 1 percent [1]. If this relationship is accepted as being correct, then a precision measurement of $g(\Delta Z)_0/\nu$ in fact determines the product hN, since M^*_{He} is known very accurately. Therefore, this measurement can be used as a stochastic input in the determination of the fundamental constants [2].

We have taken the "final" 12 stochastic inputs of Taylor et al. [2] plus possible values X and standard deviations σ_x of $g(\Delta Z)_0/\nu$ as an additional input and used their least squares adjustment procedure to determine changes in the values of α^{-1}, e, N and the electrical conversion factor K. The sensitivity (logarithmic derivative) of the values of these fundamental constants to possible values of $g(\Delta Z)_0/\nu$ denoted by X, are shown in figure 1 as a function of assumed standard deviations σ_x. It is apparent that N and e will be changed the most by a value of $g(\Delta Z)_0/\nu$ different from that given by the present values of the fundamental constants and α^{-1} will change the least. Also one can see that reducing σ_x below 2 ppm does not further increase the sensitivity of the values of α^{-1}, e, K, and N to the values of $g(\Delta Z)_0/\nu$.

* Present Address: Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556.

We have also made these same computations (a) using the 12 "final" inputs minus the three μ_p/μ_n measurements and (b) using the 12 "final" with the new measurement[3] of $2e/h$ replacing the old. The limiting values of the sensitivities at small values of σ_x are given in the table 1.

One observes from this table that the sensitivities

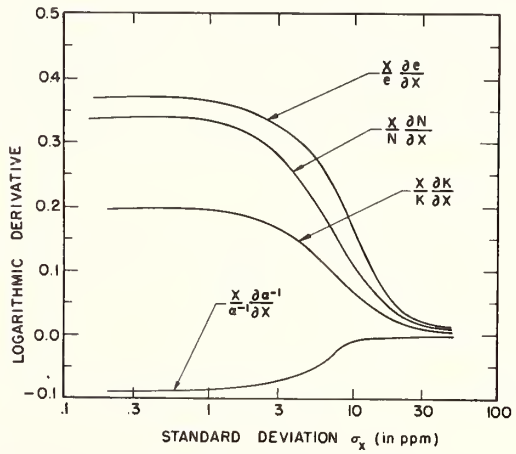


FIGURE 1. *Sensitivity (logarithmic derivative) of the value of the electronic charge, the fine structure constant, Avogadro's number and the electrical conversion factor to a measurement of $g(\Delta Z)_0/\nu$ (equals X) for liquid helium.*
The "final" 12 stochastic inputs of Taylor, Parker, and Langenberg were used.

are not changed significantly by these three different choices of stochastic inputs.

In conclusion, we find that an accurate measurement of the ac Josephson effect in liquid helium would be quite important in a new determination of the fundamental constants with the electronic charge e and Avagadro's number N being affected the most.

TABLE 1. *Limiting values of logarithmic derivatives for different sets of stochastic inputs*
(X Equals $g(\Delta Z)_0/\nu$)

Set of Stoch. Inputs	$(X/\alpha^{-1})/(\partial\alpha^{-1}/\partial X)$	$(X/e)(\partial e/\partial X)$	$(X/K)(\partial K/\partial X)$	$(X/N)(\partial N/\partial X)$
I. "final" 12 of Taylor et al. [2]	−0.087	0.37	0.20	0.33
II. Set I minus the three μ_p/μ_n measurements	−0.042	0.29	0.20	0.46
III. Set I with new measurement [3] of $2e/h$ replacing the old	−0.062	0.29	0.22	0.49

Acknowledgments

This work was supported in part by the National Science Foundation. The computations were performed at the University of Rochester Computing Center which is supported in part by the National Science Foundation.

References

[1] Khorana, B. M., and Douglass, D. H., Jr., Proceedings of Eleventh International Conference on Low Temperature Physics, p. 169 (1968).
[2] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Reviews of Modern Physics **41** (1969).
[3] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).

DISCUSSION

D. N. LANGENBERG: Would you care to speculate on the ultimate precision or accuracy that you think you might get with infinite amounts of money and time?

B. M. KHORANA: As you can see, what is involved in this measurement is a measurement of the height difference ΔZ, the frequency ν, and the acceleration g. You will notice that there are no electrical measurements or conversion factors involved. g can be measured to much better than a part per million and that ΔZ distance of something like a centimeter can be interferometrically measured to something like a part per million, which is about 100 Å. The frequency of course is no problem, so that eventually it should be possible to measure this product hN to—I would say it would be possible to measure this to something like a few parts per million, definitely. The main problem is going to be mechanical vibrations which produce a certain amount of noise in the system. We are in the process of taking care of that by mounting everything on isolation tables and things like that.

X-RAYS

Problems in Relative and Absolute Measurement of X-Ray Wavelengths

J. A. Bearden

The Johns Hopkins University, Baltimore, Md. 21218

The technique of establishing standard x-ray wavelengths should provide not only the best values for present and new measurements, but also a means of evaluating the several thousand x-ray lines previously measured. The quality of crystals and index of refraction affects wavelength ratios. Adjustment errors of a double crystal spectrometer produce important wavelength errors. Unexplained discrepancies in recent measurements are discussed. Absolute wavelengths established by a combination of Avogadro's number with crystal data, ruled grating measurements, and the use of x-ray interferometry are discussed.

Key words: Avogadro's number; double crystal spectrometer; index of refraction; standard x-ray wavelengths.

1. Introduction

Most x-ray spectroscopic measurements of relative wavelengths depend on Bragg reflection from crystals. The index of refraction is of major importance in the measurement of wavelengths greater than 2 Å; for shorter wavelengths flat polished crystals must be used if accuracies of the order of a few parts per million are to be attained. The quality of the crystals affects wavelength ratios. Adjustment errors of a double crystal spectrometer cause important wavelength errors. Some unexplained discrepancies in our recent precision measurements are discussed.

Absolute wavelengths may be determined either through a combination of Avogadro's number with

crystal data or by grating measurements or by procedures employing x-ray interferometry. If the recent major change in the Avogadro number is accepted, variations in and lack of precise measurements of atomic abundances contribute the remaining largest uncertainty in the evaluation of a crystal grating from its molecular weight and density. Analysis of the concave grating applied to x-ray measurements raises serious questions as to its accuracy. Plane grating measurements of the AlKα line may require a correction for Fresnel diffraction, but this and the ruled grating experimental errors are probably smaller than those encountered in crystal measurements of this wavelength.

The implications of the above factors in the selection of the W K$\alpha_1 = 0.2090100$ Å* as a primary x-ray wavelength standard and the evaluation of the secondary standards are discussed with reference to an extensive redetermination of these wavelengths with nine selected crystals of silicon, calcite, and quartz.

2. Refraction and Spectrometer Errors

The simplicity of the Bragg law and its seemingly wide applicability to both a large range of x-ray wavelengths and a variety of crystals has perhaps led to the misconception that the limitation in the precision of x-ray wavelengths and crystal dimensions is primarily due to the inaccuracies in the divided circle of the spectrometer. Several years ago, in an attempt to overcome this limitation, we developed an angle-measuring interferometer capable of 0.01 second accuracy. It has been used both for direct calibration of the divided circle rulings used in measurements of the Bragg angle θ and also for the direct measurement of angles less than 30°. Over the past five years in our attempt to push the wavelength error to less than 1 ppm (part per million), it has become increasingly clear that we had previously assumed too great perfection of crystals and also too naive an application of the index of refraction correction to the Bragg law.

The following are three equivalent forms of the Bragg law for the case of surface reflection

$$n\lambda = 2d_\infty[1-(\delta/\sin^2\theta_n)]\sin\theta_n \qquad (1)$$

$$n\lambda = 2d_\infty[1-(4d^2/n^2)\cdot(\delta/\lambda^2)]\sin\theta_n \qquad (2)$$

$$n\lambda = 2d_\infty(1-\delta-\delta\cot^2\theta_n)\sin\theta_n \qquad (3)$$

where $\delta = 1-\mu$, μ is the index of refraction, θ_n the Bragg angle of incidence for peak intensity in the n-th order, λ the wavelength, and d the grating constant. In eq (3) the first δ represents the fractional change in d_∞ due to the change of wavelength of the x ray inside the crystal, while the second term, $\delta\cot^2\theta$, is due to the change of direction of the incident beam as it enters the crystal. The exit angle is of no importance for the Bragg type of single crystal spectrometer or the double-crystal spectrometer.

Two examples will indicate the problem of correcting wavelengths for the index of refraction. In the case of the Al Kα line diffracted by quartz in the first order [1], δ is of the order of 247×10^{-6} with a probable error of a few percent. The $\delta\cot^2\theta_n$ term is 2.5×10^{-6}. The crystal wavelength of the Al Kα was determined by measuring the ratio of its diffraction angle in the first order with quartz $(10\bar{1}0)$ to that of the Cu Kα_1 in the third order, for which $\delta = 8.55\times10^{-6}$. The refractive corrections do not cancel, and the major error in the wavelength ratio is due to the inaccuracy in the value of δ used.

A further serious problem arose when an attempt was made to check the quartz results with an ADP (101) crystal. No experimental values of δ/λ^2 are available for ADP; hence its value was calculated from theory. At the Cu Kα and Al Kα wavelength the values of δ/λ^2 are 2.56 and 2.55×10^{-6} respectively. Use of the ADP gave an Al Kα wavelength 26 ppm lower than with quartz. The statistical probable error for each was less than 1 ppm; thus the results were completely incompatible. The shape of the Cu Kα_1 line recorded in the first and third order was quite different, indicating that the ADP crystal may not be suitable for the precise measurement of x-ray wavelengths.

In the measurement [3] of the Ag Kα_1 line with a selected calcite crystal, it was observed that a surface grinding and etching procedure could change the observed wavelength by 20 ppm. A HNO_3-detergent etch appeared to eliminate the error as compared to measurements in transmission. Recently we observed errors of 10 to 15 ppm in the wavelength of the Cr Kα_2 line measured in the first order with a quartz $(10\bar{1}1)$ plane and with two very perfect silicon crystals cut to expose the (111) plane. Measurements with other crystals of different types and with planes which gave larger Bragg angles yielded consistently longer wavelengths. Measurements were made both with a single crystal spectrometer and with a double. All crystals had been fine ground (5 mieron abrasive) parallel to within 10 to 20 s of the atomic planes and etched to maximum resolving power.

Subsequently the (111) silicons and the quartz were optically polished, the polishing procedure being continued for two to three hours after a good optical surface was obtained. The resolving power was then approximately equal to that of the etched crystals. The wavelength of the Cr Kα_2 was remeasured both in the single and double-crystal arrangement and the observed wavelength increased to that of the other crystals. A second order measurement with the ground quartz also gave a wavelength consistent with the other crystals. The first order angle for these crystals is about 20°, and the corrections for the index of refraction 125 ppm. The effective refraction angle for a ground surface is probably greater than θ and thus may account for the observed shorter wavelengths.

Although there was no clear evidence that grinding and etching procedures had influenced the measured wavelengths with the other crystals, they were all deep polished as a precautionary measure. The use

of crystals of small grating constant and high diffraction orders further reduces the refraction correction and hence its error, as indicated by eq (2).

Another experimental observation is that in transmission measurements with the W $K\alpha_1$ line, differences in wavelength of one to three ppm are observed when the crystal is rotated 180° front to back. In these cases the crystal surfaces (front and rear) are perpendicular to the atomic planes to within 10 to 20 s and both surfaces ground and etched alike. Polished crystals have not been tried. In this transmission geometry the index of refraction effect should be zero. Small crystal imperfections may account for the discrepancy, but are too small for present detection techniques. The perfection of the crystals is such that in surface reflection of short wavelengths (W $K\alpha_1$) more than 25 percent of the x rays diffracted by a crystal enter the crystal and can travel parallel to the surface as much as 30 mm before emerging from the end of the crystal in a beam parallel to the normally diffracted beam. In most of our measurements over the past five years we have covered the ends of the crystals with lead stops to exclude this part of the diffracted beam from the recorded intensities.

DuMond and Hoyt [4] pointed out that in double crystal spectrometers with the first crystal fixed, a "geometric" intensity curve, due to slit widths combined with any inhomogeneity in the focal spot or in the reflecting surface of the crystals, will be superimposed on, and hence distort, the spectral line under study. A test was recently made on the magnitude of the possible error that might occur in a precision wavelength measurement due to these effects. A stop was placed between the crystals which intercepted one-half the normally observed intensity of a diffracted W $K\alpha_1$ line. The apparent wavelength of the x-ray beam which passed the stop was measured. The stop was then inserted on the opposite side of the x-ray beam and the wavelength remeasured. The difference in the two measurements was 22 ppm. A similar difference was observed when the stop was placed between the first crystal and focal spot. These observations indicate the extreme requirements in accuracy of adjusting stops and slits to exclude background without affecting the wavelength of the measured line. Simultaneous rotation of both crystals would avoid this difficulty, but no spectrometer design of this type for absolute measurements is known to the author. The above problem is almost nonexistent with a single-crystal Bragg spectrometer.

The observations described above clearly indicate that with our presently available crystals and knowledge of crystal dynamics, the establishment of relative wavelengths to better than 1 ppm is beset with numerous problems. The only practical solution at present seems to be to rely on use of several different crystals and planes in establishing the ratios of secondary standards to a primary standard. The establishment of the precise grating constant of a crystal, while of great importance for other purposes, will aid the present wavelength situation only as better crystals are available and more clearly understood.

3. Scale of X-Ray Wavelengths

It is well established that we have had two mutually incompatible working standards for x-ray wavelengths, viz., the wavelength of the Mo $K\alpha_1$ line = 707.831 xu, and that of the Cu $K\alpha_1$ = 1537.400 xu. Recent measurements [5] by the author indicate that the resulting ratio is in error by approximately 20 ppm. In addition many x-ray lines have been measured with calcite crystals of unknown grating constant but with an assumed $d_1 = 3029.04$ xu. The rather general practice in the past forty years of using either of the two wavelengths or the calcite $d_1 = 3029.04$ xu for measuring x-ray wavelengths and crystal dimensions and the confusion as to the relation of these to the angstrom clearly indicated to the author [5] a few years ago the necessity for defining a new and simple standard reference wavelength together with suitable substandards. The W $K\alpha_1$ wavelength was taken as 0.2090100 Å* (a defined value without probable error) and chosen as the primary standard; the Ag $K\alpha_1$, Mo $K\alpha_1$, Cu $K\alpha_1$ and the Cr $K\alpha_2$ lines were selected as secondary standards [5].

The numerical value of the primary standard in Å* was chosen as nearly equal to the value in angstroms as experimental data permitted at that time. New measurements and analysis of atomic constants [6] indicate that these values differ by 19 ppm. However, new measurements of the W $K\alpha_1$ to Cu $K\alpha_1$ ratio, corrected for some of the refraction errors discussed above, lower this to 11 ppm, which is about twice the original estimated probable error of 5 ppm. This does *not* mean that the wavelength standard must be redefined. Like other relative standards, its use should be continued until a new primary standard of much higher precision is established. In the meantime all new precision measurements can be made on a standard such that they can accurately be reevaluated on some new standard. Since few users of x-ray wavelengths require a precision greater than 10 to 20 ppm, the values in Å* can be considered equal to those in Å for most purposes.

The experimental absolute value of the W $K\alpha_1$ line = 0.2090100 Å was obtained [7, 8] from the density of silicon and calcite crystals, the x-ray diffraction angles for the Cu $K\alpha_1$, the indices of refraction, and the 1965 non-x-ray value [9] of Avogadro's number $N = 6.02252\ (\pm 9) \times 10^{26}$ k mol^{-1}. A large part of the uncertainty in this value arose from the use of the above recommended value of Avogadro's number, as opposed to the presently recognized lower value $N = 6.02217\ (\pm 4)$. An additional significant error is due to the uncertainty in the relative atomic abundance of the isotopes of silicon [7].

New ruled grating measurements of the Al $K\alpha$ line by Henins [1] (also reported at this Conference)

give

$$\lambda(\text{Al K}\alpha) = 8.34034\ (\pm 5)\ \text{Å}.$$

His value of the quartz crystal wavelength based on the Å* standard above = 8.34027 Å*; thus he obtains $\Lambda^* \equiv \lambda(\text{Å})/\lambda(\text{Å}^*) = 1.000\,009$ Å/Å*, as compared with the value $\Lambda^* = 1.000\,019$ given by Taylor et al. [6].

4. A Remeasurement of the Primary and Secondary Standards

With the possibility that the x-ray interferometer will soon make available a high precision grating constant of silicon, it is essential that new crystal diffraction studies be made to explore the limitations of relative x-ray wavelength measurements. Three spectrometers have been used: (1) a single-crystal Bragg type with slits 0.025 mm wide separated 500 mm; (2) a double-crystal instrument for tungsten radiation with 180 cm separation between tube and detector to reduce scattering; (3) a double-crystal with small separation of source crystals and detector to record wavelengths to 2.29 Å*.

A large ingot of high perfection silicon was cut to furnish a pair of each of the (220), (111), and (400) planes, and one (111) pair for transmission. The (220) was also used in transmission. A separate pair of excellent (220) silicon was available from an older crystal. Three pairs of our best calcite crystals were included, obtained from the following sources: (1) Iceland; (2) Argentina, S.A.; (3) National Bureau of Standards, selected, but origin unknown. Two quartz crystals were used, one synthetic $(10\bar{1}0)$ and the other natural quartz $(10\bar{1}1)$, most probably Brazilian. The crystals were first ground parallel to the desired atomic plane within 10 to 20 s of arc. The final grinding was done on very flat laps with 5 micron aluminum oxide. They were then etched to about twice the time to reach maximum resolving power as indicated by (1−1) measurements. In later measurements the surfaces of all crystals were reground with 2 micron material and then optically polished (deep polish). The resolving powers were slightly less than the etched crystals.

The spectrometer circle was read by four microscopes in which the cross hairs had been replaced by a slit and photocell. The circle has been calibrated for each 6°40′ interval such that every 26°40′ interval ending on a 6°40′ line is known within 0.05 s. Nine calibrations have been made at intervals of several months. In addition an angle-measuring interferometer has been used to further subdivide the calibrated intervals and every circle line used in measuring a Bragg angle has been calibrated at least twice. Thus the precision of our wavelength measurements is not limited by angular errors due to the spectrometer.

Temperature of the spectrometer and crystals was maintained to within a few hundredths of a degree of 25.00 °C.

The x-ray intensity as a function of crystal angle is recorded with NaI (Th) detectors and the usual electronics on a model 35 teletype, together with the photocell currents (converted to pulse rate) obtained in reading the rulings on the circle. A computer analyzes this data, determining the peak of the x-ray intensity curves and the minimum in the photocell curves for each of the four microscopes; it then computes the x-ray wavelength.

Most of the W $K\alpha_1$ measurements were made in the 3rd to 6th orders where the total correction for refraction is less than 10 ppm. The tungsten Bragg angles from polished crystals are from 1 to 2 ppm greater than those with ground and etched surfaces. In the cases of Ag, Mo, and Cu in high orders the differences were less. The Cr measurements were made mostly in first order, a few in second, and only two in the third order. For Cr large differenees (10 to 15 ppm) were obtained between ground and polished crystals for the silicon (111) and quartz $(10\bar{1}0)$ planes, as noted above.

With the W $K\alpha_1$ as primary standard the new measurements indicate an increase of approximately 10 ppm in the wavelengths of the Ag, Mo, and Cu $K\alpha_1$ lines, but only 2 to 3 ppm for the Cr $K\alpha_2$.

5. References

[1] Henins, A., Thesis, Johns Hopkins University (1969).
[2] Schnopper, H. W., J. Appl. Phys. **36**, 1415, 1423 (1965)
[3] Bearden, J. A., Bull. Am. Phys. Soc. **7**, 339 (1962).
[4] DuMond, J. W. M., and Hoyt, Archer, Phys. Rev. **36**, 1702 (1930).
[5] Bearden, J. A., Phys. Rev. **137**, B455 (1965).
[6] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 2 (1969).
[7] Henins, I., and Bearden, J. A., Phys. Rev. **135**, A890 (1964).
[8] Bearden, J. A., Phys. Rev. **137**, B181 (1965).
[9] Cohen, E. R., and DuMond, J. W. M., Rev. Mod. Phys **37**, 537 (1965).

DISCUSSION

M. Hart: If one works in symmetric transmission instead of in reflection, one finds there is no refractive index correction to the Bragg angle. This is quite feasible with existing crystals like quartz and silicon. Even with a wavelength of about 4 Å you can still cut crystals thin enough. There is no absorption to worry about. Why not eliminate the correction? It's far easier.

J. A. Bearden: I mentioned the apparent zero effect of refraction at normal incidence but when the crystal is rotated 180 degrees and one observes a difference in wavelength, then what's the answer? Also in reflection, if one includes in the intensity the beam which emerges from the end of the crystal, again one observed a different wavelength. Could this not also be a factor in normal transmission measurements?

M. Hart: You still have reflection on the first crystal, while if you have both crystals in symmetric transmission there's no correction to worry about. This is a far better approach to eliminate your systematic error rather than try to make corrections for them.

Ruled Grating Measurements of the AlK$\alpha_{1,2}$ Wavelength

Albert Henins[1]

The Johns Hopkins University, Baltimore, Md. 21218

An absolute wavelength measurement of the AlK$\alpha_{1,2}$ unresolved doublet was made by use of a plane grating of nearly ideal surface contour. The same wavelength was also measured by a crystal spectrometer with respect to the CuKα_1 line to obtain the conversion factor Λ^*.

Preliminary analysis of the data yields the wavelength of the AlK$\alpha_{1,2}$ as 8.34034 Å and the conversion factor $\Lambda^* = 1.000\,009$ Å/Å$^* \pm 7$ ppm P.E.

Key words: Absolute x-ray wavelength; ruled grating; x-ray conversion constant; x-ray spectroscopy.

1. Introduction

In x-ray spectroscopy, as in many other fields of physics, it is usually possible to make relative measurements with greater accuracy than absolute ones. Since most x-ray wavelength measurements are made by crystal diffraction, the resultant wavelengths depend on a knowledge (or a definition) of the grating constant of the crystal employed or the crystal grating constant in terms of a known (or a defined) wavelength. In neither case is the result obtained in absolute terms. Consequently, precision measurements of x-ray wavelengths and crystal parameters are best known on a relative, rather than an absolute scale.

There are three methods for the determination of absolute wavelengths of x-rays. The most direct way is by use of ruled gratings. The second is by use of the Bragg equation which determines the unit cell dimensions of a crystal from its atomic weight, density, and Avogadro's number [1, 2]. The third and most promising is the use of optical and x-ray interferometer technique [3] which determines the unit cell dimension in terms of optical wavelength. The present measurements made use of a ruled grating in which new techniques were employed to extend the precision of this method to the limit presently possible with this method.

2. Experimental Arrangement

In the initial work an attempt was made to use the angle measuring interferometer [4] to improve on Bearden's [5] measurements where the grating is placed between the crystals of a double crystal spectrometer. Crystal imperfections and dispersion [6] made this method impractical. A second method was to measure the incident and diffracted angles with interferometer. At the small angles involved it was found that the stringent centering requirements were unattainable. The method finally adopted was a refinement of that used by Bearden [7] in 1931.

The method employed is basically shown in figure 1. A very narrow slit S and the narrow line focal spot form a highly collimated x-ray beam with a half-width approximately 10 sec of arc. This beam falls on the ruled grating G of grating constant d, which is supported on the axis of a precision spectrometer which is also used to measure the crystal diffraction angles. P_1 and P_2 are photographic plates (Ilford Q2)

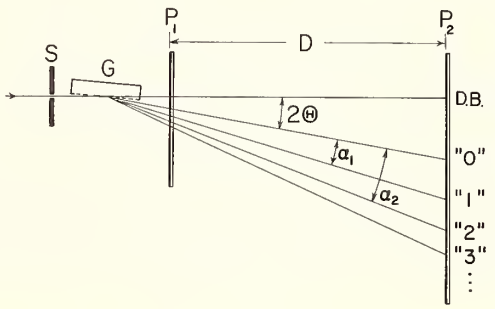


FIGURE 1. *Experimental arrangement.*
Incident x-ray beam upon grating G is collimated by a line source (not shown) and slit S. P_1 and P_2 are photographic plates; D is separation of plates; D.B., "0", "1", "2", . . ., are respectively the positions of direct beam, zero order reflection and diffracted orders upon the photographic plates.

upon which the direct beam (D.B.) and the various diffracted orders are recorded. These plates are held by special holders to place them a precisely measured distance D apart as well as to make them parallel to each other and perpendicular to the incident beam. In order to determine a wavelength with this experimental arrangement, one has to know the grating constant d, the separation of the plates D, and the distances between the various lines on the plates.

Since ruled gratings must be used at grazing incidence for x-ray spectroscopy, the usual grating

[1] Present address: National Bureau of Standards, Washington, D. C. 20234.

equation is expressed in terms of angles measured from the surface rather than the normal of the grating. For purposes of calculation it is more convenient to express the wavelength as a product of sines rather than as a difference of cosines, hence

$$n\lambda = 2d \sin\tfrac{1}{2}(2\theta+\alpha_n) \sin\tfrac{1}{2}(\alpha_n) \qquad (1)$$

where θ is the angle between the surface of the grating and the direct or ineident beam, and α_n is the angle between the reflected beam (zero order) and the diffracted beam in the nth order. Hence $2\theta+\alpha_n$ is the angle between the direct (incident) beam and a diffracted beam.

The angle measuring interferometer was used as a monitor on the angular position of the grating. Increase in accuracy over the 1931 measurements was dependent upon using a better grating, working with larger diffraction angles by choosing a longer wavelength, and improving the accuracy of measurement. The choice of a longer wavelength also had the desirable effect of increasing the critical angle for the grating, which allowed one to work in a region well away from the critical angle.

Normally all spectroscopy at these wavelengths is done in a vacuum due to the high absorbtion by air (approximately 80%/cm). The use of a hydrogen atmosphere for the beams not only provided a relatively absorbtion-free path, but also avoided the serious mechanical complications that arise when one attempts to perform such experiments in a vacuum.

Equation (1) was derived by considering parallel beams incident upon and emerging from the grating, i.e., Fraunhofer diffraction. Due to the finite source to grating and grating to photographic plate distances, these beams cannot be truly parallel, but must exhibit a certain amount of divergence. This means that Fresnel diffraction effects will be present and may affect the results. For the experimental geometry used application of theoretical calculations of Thomsen [8] for the zero order gave a maximum wavelength shift of less than 2 ppm. The shifts for the diffracted orders have not yet been calculated, but it is felt that their effect would tend to cancel the effect by the zero order and hence no corrections due to Fresnel effects will be considered.

3. Grating

In the wavelength region below 500 Å, diffraction gratings are used at grazing incidence and have been much less efficient than when used at longer wavelengths. This was demonstrated by the early failures of Kirkpatrick et al. [9] in trying to repeat Tyrén's [10] measurements. Franks [11] has shown that this inefficiency is the result of inevitable irregularities in the groove profile caused by the breakup of the glass surface or by the debris thrown up in the process of ruling a metal or glass surface.

When a plane grating is used for absolute wavelength determination, as was done by Bearden [12], the profile of the land or flat is of great importance. If this flat is not a flat but exhibits regions with blaze, then it is quite possible that the spectrum obtained will exhibit abnormal behavior of the type found by Fabre [13] or Crisp [14] where the spectrum obtained depended on the direction from which the grating was illuminated. Thus the conventional process of ruling gratings is unsuitable for the production of x-ray gratings and other methods should be sought for producing an unblemished optical surface interrupted by equally spaced nonreflecting grooves. This idealized surface contour is desired to match the assumptions used in theoretical calculations.

To produce a grating with the desired properties a new approach to the ruling of gratings was developed; i.e., the rulings were etched into the grating blank. The glass grating blank was first polished by the normal optician's procedures used to produce the smoothest flat surfaces obtainable. The smoothness of the blank was determined by examination of replications of the surface in an electron microscope. When no further improvement could be obtained, the polishing was stopped and the blank was coated (by sputtering) with an etchant resistant coating. This coating consisted of a gold layer some hundreds of angstroms thick, with a thin molybdenum undercoating to improve the film's adherence properties. The blank was then ruled (7500 lines/inch) with a sharp, symmetrically shaped tool. The ruling process was used only to cut through the protective film and to open lines along which the etchant, a 2 percent solution of hydrofluoric acid, could act upon the grating blank and produce grooves with relatively steep sides. After etching, the protective coating was removed and the grating surface was recoated with platinum. Sample gratings with 15,000 lines/inch have also been produced by this technique.

4. Grating Measurement

The grating constant, d, for the area used was determined by use of an optical interferometer; hence the grating constant was obtained in terms of the (5641 Å) green line of mercury. The separation of the surfaces against which the photographic plates were placed, the plate separation D, was obtained by comparing it to a nearly equal distance set by a gauge block [15] caliper. The distances along the photographic plates were measured by a Gaertner comparator with a photoelectric display [16] whose screw was calibrated by use of an interferometrically calibrated glass scale.

The diffraction angles and the angles of incidence were determined from measurements of the separation of the various lines along the plates (direct beam, zero order, and diffracted orders), and the knowledge of the plate separation D. The wavelength was then calculated by use of eq (1).

The results for the various orders are shown in table 1. The weighted average wavelength (grating) from 72 plates of up to six orders each was

$$8.34034 \pm 0.00005 \text{ A P.E. or } \pm 6 \text{ ppm.}$$

Table 1. *Grating measurements*

Order	Average wavelength, Å	Average deviation	Number of plates	Weight
1	8.34031	±0.00043	67	2
2	35	30	71	4
3	23	25	72	4
4	52	27	71	4
5	25	31	54	3
6	38	28	38	2

As the difference between values obtained for the six orders exceeded those expected from estimated errors in the length measurements involved, the final assigned probable error is taken as the average deviation from the mean of the results for the various orders.

5. Crystal Measurements

The wavelength of the AlK$\alpha_{1,2}$ line was determined by measuring the ratio of its diffraction angle to that of CuKα_1 with a quartz ($10\bar{1}0$). In this measurement the crystal replaced the ruled grating with no changes in the x-ray tube or collimation. This procedure minimizes the criticism of Sandström [17] concerning the probable chemical changes in the x-ray source. Flow counters were used for detectors, and hydrogen was employed for all beam paths to minimize absorbtion. Since the grating measurements yield the centroid wavelength of the AlK$\alpha_{1,2}$ doublet the same feature was also measured with the crystal; however, for the CuKα_1 the peak wavelength was determined.

The wavelength ratio obtained is strongly dependent upon the value of the index of refraction correction used at the AlK$\alpha_{1,2}$ wavelength since this measurement is done in the first order, whereas the CuKα_1 is measured in the third order. This index can be either calculated theoretically or determined experimentally [18]. The values of δ/λ^2 (in Å) for quartz as determined theoretically and experimentally are respectively 3.615×10^{-6} and 3.64×10^{-6} at 1.5 A, and 3.39×10^{-6} and 3.47×10^{-6} at 8.3 Å.

The crystal wavelength measurements were expressed in $\lambda/2d_\infty$ rather than wavelength to emphasize the fact that the measured ratio will depend on the index of refraction correction and not on the exact value of the grating constant. Using the theoretically obtained values of δ/λ^2 eighteen data runs for the peak value of CuKα_1 gave

$$\lambda/2d = (181022.7\pm0.12 \text{ a.d.})\times10^{-6}$$

and twenty runs for the centroid of the AlK$\alpha_{1,2}$ measured in the first order gave

$$\lambda/2d = (980020.1\pm0.8 \text{ a.d.})\times10^{-6}.$$

Due to the nonlinear dispersion at large diffraction angles, the centroid of a diffraction profile calculated on an angular scale will differ from that obtained on a wavelength scale. For the AlK$\alpha_{1,2}$ measurement with quartz ($10\bar{1}0$) crystal planes this correction in 3 ppm [19] and has been applied to the above measurements.

The wavelength ratio of the AlK$\alpha_{1,2}$ to the CuK$\alpha_{1,2}$ then is

$$5.413797\pm1 \text{ ppm a.d.}$$

however, if the experimentally determined values of the index of refraction are used, then this ratio becomes

$$5.413767\pm1 \text{ ppm a.d.}$$

or 5.6 ppm smaller.

In view of the index of refraction problems, crystal measurements of the wavelength ratio were also made using ADP (101) crystal planes. The measurement accuracy was the same as that with the quartz crystal. The theoretically calculated values of δ/λ^2 for ADP are 2.563×10^{-6} at 1.5 Å and 2.546×10^{-6} at 8.3 Å. Using these values for the index of refractions, the wavelength ratio was 5.413672 which differs from the quartz crystal value by 20 ppm. An effort was made to experimentally determine the value of the index of refraction correction for ADP at the copper wavelength by the method of ratio of orders, using the first and third orders. Here it was found that the particular crystal used would give a highly unsymmetric diffraction profile in the first order, hence casting doubt on any values obtained with the crystal.

The ADP crystal surfaces are known to deteriorate as they react with water vapor in air [20]. Such surface deterioration should have more effect on the measurement of the aluminum wavelength at 8.3 Å than for copper at 1.5 A.[2]

The quartz crystal has also been used for a whole range of wavelength measurements from 2.2 Å to 0.2 Å and yielded wavelength ratios in excellent agreement with those obtained with silicon crystals. Hence in view of the problems with the ADP crystal it is considered best to ignore the ADP result and depend solely on the quartz measurement.

6. Conclusion

As the individual crystal measurements have probable errors of less than 1 ppm, the index of refraction corrections were the dominant error in the wavelength ratio. The average of the two wavelengh ratios as obtained by using the theoretical and experimental values for the index of refraction was taken as the final value for the ratio with a probable error of half the difference of the two values, i.e.:

$$5.413782\pm(2.8 \text{ ppm}).$$

Hence the wavelength of the AlK$\alpha_{1,2}$ measured with

[2] Recently I have learned from R. D. Deslattes that he had used ADP crystals for wavelength measurements and had found unexplained inconsistencies among the values of the wavelength ratios obtained by different ADP crystals.

respect to the CuKα_1 ($\lambda = 1.540562$ A*$\pm$1.2 ppm) [21] is

8.340267 A*$\pm$3.1 ppm.

The ratio of the wavelengths as measured by the grating to that as measured by the crystal then is

$\Lambda^* = 1.000\ 0087$ Å/Å*$\pm$7 ppm P.E.

If the wavelength of the copper line is taken as 1.537400 kxu, then

$\Lambda = 1.002\ 0655$ Å/kxu$\pm$7 ppm P.E.

This can be compared to 1.002 0764 Å/kxu$\pm$5.3 ppm the value given by Taylor et al. [22].

Note added in proof: When Fresnel effects for the diffracted orders are also considered the resulting calculations show a correction of about 1 ppm. Since this correction is small compared to the experimental errors involved it will be omitted and the results left as stated.

7. References

[1] Henins, I., J. Res. Nat. Bur. Stand. (U.S.) **68**, 529 (1964).
[2] Henins, I., and Bearden, J. A., Phys. Rev. **135**, 890 (1964).
[3] Deslattes, R. D., Appl. Phys. Letters **15**, 386 (1969).
[4] Marzolf, J. G., Rev. Sci. Instr. **35**, 1212 (1964).
[5] Bearden, J. A., Phys. Rev. (L) **47**, 883 (1935); Phys. Rev. **48**, 385 (1935).
[6] Sauder, W. C., Dissertation (The Johns Hopkins University, Baltimore, Md., 1963), p. 122; J. Appl. Phys. **37**, 1495 (1965).
[7] Bearden, J. A., Phys. Rev. **37**, 1210 (1931).
[8] Thomsen, J. S., Private communication.
[9] Kirkpatrick, H. A., DuMond, J. W. M., and Cohn, E. R., "Remeasurement of the Conversion Constant, Λ," Proceedings of the Third International Conference on Atomic Masses, Baker, R. C., Ed. (University of Manitoba Press, Winnipeg, 1967), p. 347.
[10] Tyren, F., Z. Physik **109**, 722 (1938), and Nova Acta Reg. Sci. Upsal. (IV) **12**, No. 1 (1940).
[11] Franks, A., in X-ray Optics and X-ray Microanalysis, Pattee, H. H., Cosslett, V. E., and Engstrom, A., Eds. (Academic Press, New York, 1963), p. 199.
[12] Bearden, J. A., Proc. Natl. Acad. Sci. U.S. **15**, 528 (1929).
[13] Fabre, D., Optica Acta **15**, 171 (1968).
[14] Crisp, R. S., Optica Acta **8**, 137 (1961).
[15] I am indebted to Mr. Philip A. Moss of the Martin-Marietta Corporation of Baltimore, Md., for the loan of the N.B.S. calibrated gauge blocks and the Scheffield Accutron.
[16] Dieke, G. H., Dimock, D., and Crosswhite, H. M., J. Opt. Soc. Am. **46**, 456 (1956).
[17] Sandström, A. E., in Handbuch der Physik, Flügge, S., Ed. (Springer-Verlag, Berlin, 1957), Vol. XXX, p. 240.
[18] Larsson, A., Inaugural Dissertation, Uppsala Universitets Arsskrift (1929), pp. 103–106.
[19] Thomsen, J. S., Private communication.
[20] Deslattes, R. D., Torgesen, J. L., Paretzkin, B., and Horton, A. T., in Advances in X-ray Analysis **8**, Mueller, Mallett and Fay, Eds. (Plenum Press, New York, 1965), p. 315.
[21] Bearden, J. A., Rev. Mod. Phys. **39**, 78 (1967).
[22] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

Precision Atomic Binding Energies

A. F. Burr

New Mexico State University, Las Cruces, New Mexico

A very brief summary of the method for obtaining atomic binding energies for occupied energy levels by means of x-ray line excitation curves is given with some results of measurements performed on cobalt, nickel, lanthanum, and cerium. A very condensed description of a method for using L_{II}–L_{III} binding energy differences to obtain an estimate of the Lamb shift for heavy elements is also given.

Key words: Binding energies; excitation curves; Lamb shift; quantum electrodynamic effects in heavy elements; x-ray emission lines.

1. Binding Energies

Basic to our understanding of the nature of atoms is a knowledge of the amount of energy necessary to remove one of the several electrons which make up the atom. This energy, of an amount just necessary to remove the electron from the atom, is called the binding energy for the electron under study.

There are four major sources of information which can be used for determining the binding energy of the various electrons in each element. They are given in table 1. In practice binding energy values have been obtained mainly from the information presented by x-ray emission lines and photoelectron spectroscopy because photoelectron lines

Table 1. *Sources of information about binding energies*

1. X-ray absorption edges
2. X-ray emission lines
3. Photoelectron spectroscopy
4. Excitation curves of x-ray lines

yield the same information as x-ray absorption edges and are easier to interpret and there are not enough precision excitation curves available to contribute significantly to the information on any one element.

To obtain values for the binding energies for a given element, all the relevant information on that element is studied. Because there are usually many more emission lines, which represent binding energy differences, than the number of energy levels to be found, the procedure used is to perform a least squares analysis of all the available data with values of all the binding energies as the output of that analysis. This task has been done [1]; and a detailed analysis of the method used, and the difficulties met, given [2, 3]. Even the best known binding energy has a probable error of 0.3 eV despite the fact that some of the data which goes into determining it has a probable error of less than 0.05 eV. The reason for this state of affairs is that all binding energies are placed on an absolute scale by means of a series of measurements the weakest link of which is the measurement of a work function, an error which is common to all the binding energies. Thus in every case where absolute binding energies are not required, an improvement in precision can be obtained by using relative binding energies.

2. Excitation Curves

Binding energy information can be extracted from excitation studies of x-ray emission lines. It is clear that a given x-ray emission line cannot be emitted until a vacancy has been created in the innermost of the two levels involved in the transition which gave rise to the line. Thus, if some oversimplification is permitted, one can obtain a value for the binding energy by setting a spectrometer at the wavelength of an x-ray line and increasing the voltage across the x-ray tube until the line is observed. The voltage across the tube at the time the line is first observed will give the energy necessary to just remove the inner electron, i.e., the binding energy [4, 5]. A plot of tube voltage versus photon counting rate (an excitation curve) will illustrate the results of the experiment. One such curve is given in figure 1. It is clear that, especially when it is noted that the error in the value assigned to the binding energy of the lanthanum M_{IV} level is only 0.4 eV, there will be some difficulty in assigning an excitation voltage to the lanthanum $M_{IV}N_{III}$ line which will yield a binding energy with an error approaching that of other methods.

Indeed, the points listed in table 2 must be considered before useful binding energies can be ob-

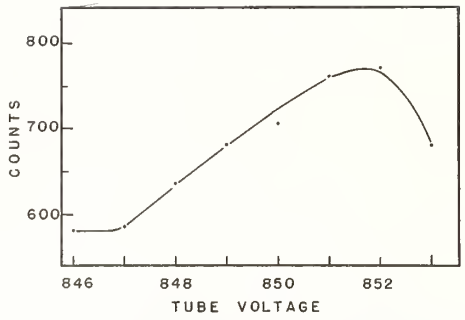


FIGURE 1. *Lanthanum* $M_{IV}N_{III}$ *excitation curve.*

tained from the excitation curve method. The first three items require a correction to the anode voltage reading before the equivalent binding energy can be obtained. The last two items affects the slope of the excitation curve and indicate that the first inflection point should be selected as the excitation voltage [6]. When one selects a line which has a narrower inner level and arranges the experimental apparatus so that, while the measurements are being taken, the filament structure is entirely at zero potential [7], much clearer exeitation curves are obtained. One such example is given in figure 2. In this case it is not hard to see that an excitation voltage could be selected with a reproducibility of the same order of magnitude as the error (0.3 eV) for the L_{III} binding energy of cobalt.

Preliminary results of experiments with some transition metal L levels and lanthanide M levels indicate that this method can be expected to yield binding energies with a probable error of about the same order of magnitude as other methods. Table 3 lists some representative levels and their binding energies as obtained by both the least squares analysis (using emission line and photoelectron data) and the excitation curve method. The numbers in the last eolumn are preliminary results which have errors somewhat larger than the probable errors listed in the preceding column.

The largest source of error comes from the work function correction and from the difficulty of assigning an exact turn-on point to the excitation eurve. The first error can be eliminated and the second reduced if absolute binding energies are not required and binding energy differences are acceptable. Table 4 lists some representative binding energy differences. The numbers in the last column are preliminary results which have probable errors

TABLE 2. *Complications to be considered when obtaining binding energies from excitation curves*

1. Power supply calibration
2. Work function of x-ray tube filament
3. Effect of the voltage drop across the filament
4. Width of the inner level

of the same order of magnitude as the values given in the preceding column. This method does not give a better defined binding energy difference than a well measured x-ray line but, if the selection rules involved prohibit a direct transition between the levels involved, the excitation curve method can contribute to the knowledge of the energy level structure of that element.

Some investigators [8] have suggested that binding energy values obtained by the excitation curve method differ systematically from the values obtained by other means. However, all the binding energy values obtained by the excitation curve method listed in table 3 agree within a small multiple of the probable errors listed with the present values. Nevertheless, because all these new values are on the low side, this point should be investigated further.

There are some x-ray lines which defy analysis in terms of an excitation curve. One such line is shown in figure 3. It is clear that no excitation curve taken at a single spectrometer setting could adequately represent the excitation characteristics of this line

TABLE 3. *Some binding energy values*

Element	Level	Binding energy (eV)[a]	Binding energy (eV)[b]
Cobalt	L_{II}	793.6±0.3	792.4
Cobalt	L_{III}	778.6±0.3	777.6
Nickel	L_{II}	871.9±0.4	871.2
Nickel	L_{III}	854.7±0.4	854.2
Cerium	M_{IV}	901.3±0.6	900.8
Cerium	M_V	883.3±0.5	883.0

[a] From reference [1].
[b] Preliminary data by the excitation curve method.

at near threshold voltages. Although the wave lengths of other lines have been known to shift [9], the lines which exhibit this phenomenon to the greatest degree are found in the M series of the lanthanides [10]. The explanation for this energy shift is given in terms of the reaction of 4*f* orbitals in the lanthanides to an inner vacancy [11], hence relatively few of the thousands of x-ray lines will be troubled with this difficulty.

3. Uses of Binding Energy Values

Binding energy values have been put to a number of practical uses. The analyst uses them to selcct transitions and elements to obtain x-rays of specific energy. They can be used to locate possible interfering lines, and to provide the wavelength of lines from hard to handle elements without having to experimentally produce these lines.

A particularly interesting use of binding energy tables is as a standard against which to compare the results of different methods of calculating the binding energy. In principle, the Hartree-Fock

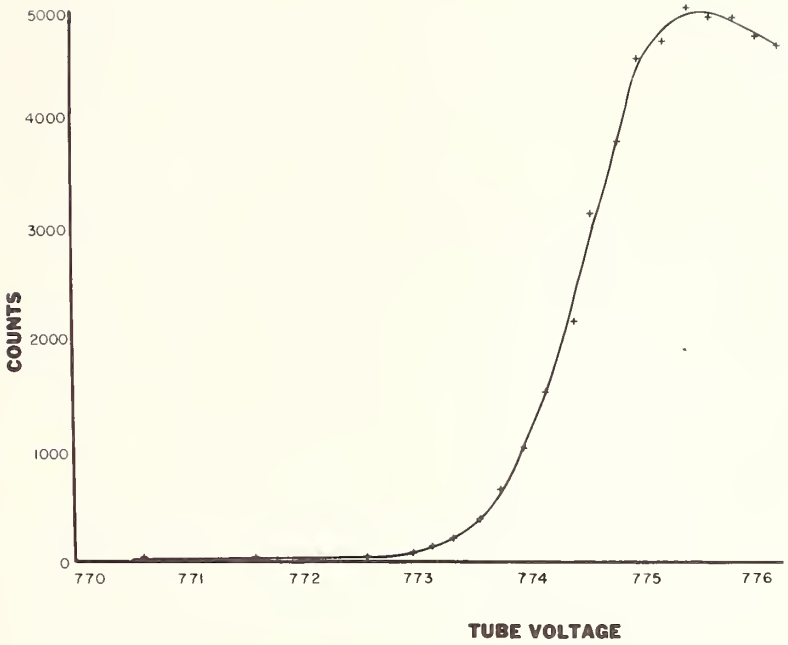


FIGURE 2. *Cobalt* Lα *excitation curve.*

method will permit the calculation of any binding energy. In practice the method is so complicated that even large computers cannot handle the problem; hence various approximations are made. The effect of these approximations can be detected by comparing the calculated binding energies with the experimental ones. In this case the precision with which the binding energy is known almost always exceeds the precision with which it can be calculated.

Binding energies, or rather binding energy differences, are also used to estimate the size of quantum electrodynamic effects in the heavy elements [12, 13]. These effects produce observable changes in the spectra of the very light elements which are in agreement with the theory. Nobody has, however, been able to calculate completely what these changes should be in heavy ($Z>70$) elements. It is clear, however, that these quantum electrodynamic changes can be detected especially in the L level binding energy differences. Table 5 lists the various effects which cause L level binding energy differences to vary from the value expected from the simple Sommerfeld theory. Some estimate of the size of these effects in the L_{II}–L_{III} splitting for the heavy elements is available for all but the Lamb shift.

TABLE 4. *Binding energy differences*

Element	Levels	B.E. Differences (eV)[a]	B.E. Differences (eV)[b]
Cobalt	L_{II}–L_{III}	15.01±0.05	14.83
Nickel	L_{II}–L_{III}	17.25±0.05	17.03
Lanthanum	M_{IV}–M_V	16.8±0.1	17.0
Cerium	M_{IV}–M_V	18.0±0.5	18.8

[a] From reference [2].
[b] Preliminary data by the excitation curve method.

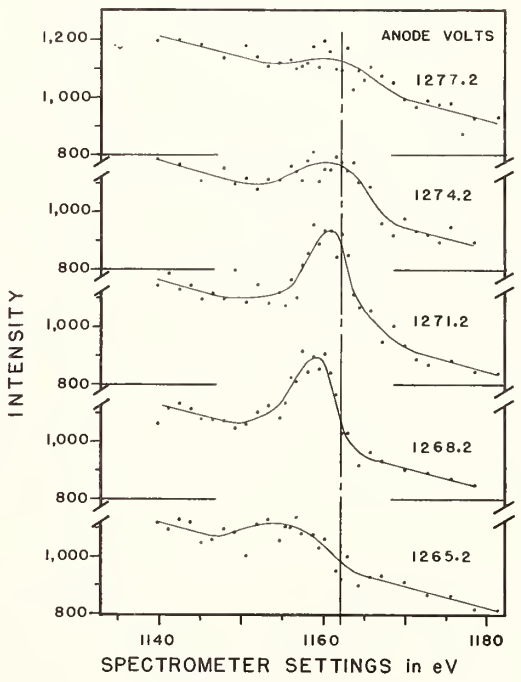


FIGURE 3. *Line profiles of the cerium* $M_{II}N_{III}$ *x-ray line taken at various anode voltages.*
The vertical line represents the spectrometer setting of the peak of the line at much higher voltages. Data taken from reference [11].

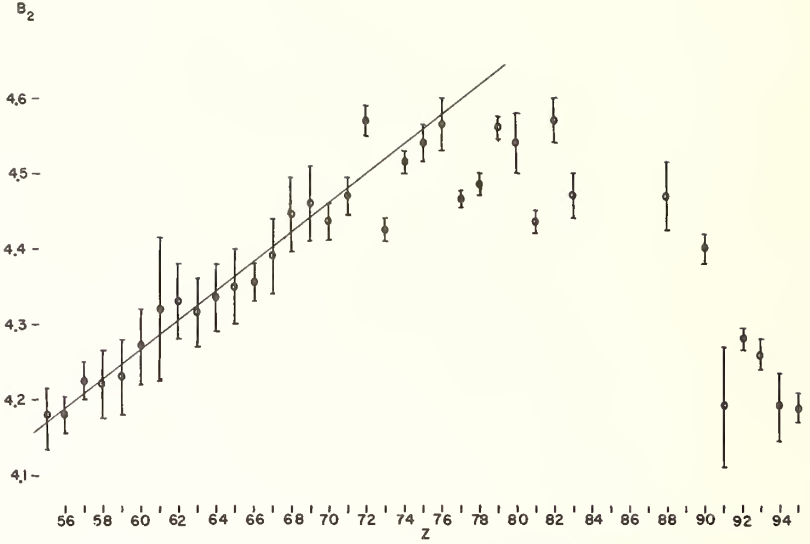


FIGURE 4. *The Z dependence of the factor* B_2 *in the* L_{II}–L_{III} *splitting formula of Christy and Keller* [14].

Experimental binding energy differences can be used to obtain this missing estimate. Briefly the method [13] is to obtain the most precise calculation of the L_{II}–L_{III} level splitting which does not include any quantum electrodynamic effects and use it to calculate the splitting which would be expected in the heavy elements. This calculated value is compared with the actual splitting. After as much as possible of the difference is assigned to known factors, the remainder is attributed to the Lamb shift.

The equation used in this case is

$$\Delta E_{L_{II}-L_{III}} = mc^2\left(\tfrac{1}{2}\{[4-(\alpha z)^2]^{1/2}-[2+2(1-(\alpha z)^2)^{1/2}]^{1/2}\} - f(\alpha z)[(\alpha z)^4/z]-0.0178[(\alpha z)^6/z]+B_2[(\alpha z)^4/z^2]\right).$$

This formula was developed by Christy and Keller [14]. Here m is the mass of the electron, c is the velocity of light, α is the fine structure constant, and z the atomic number. The terms on the first line represent the results for a hydrogenic atom with a point nucleus. The first term on the second line represents the lowest order correction factor and is obtained from the direct interaction of the K, L, and M shell electrons and the exchange interactions of the K and L shells. The function $f(\alpha z)$ is a slowly increasing function of z described in more detail by Christy and Keller [14]. The next term on the second line represents the next higher order, in αz, term and is calculated on the basis of the K shell. The last term stands for all the other neglected factors, with the constant B_2 to be calculated empirically.

A plot, as in figure 4, of B_2 versus z shows that it really cannot be considered a constant. Therefore B_2

TABLE 5. *Corrections to L level binding energy differences*

1. Finite nuclear size
2. First order anomalous moment of the electron
3. Higher order vacuum polarization
4. Higher order Lamb shift (self energy corrections)

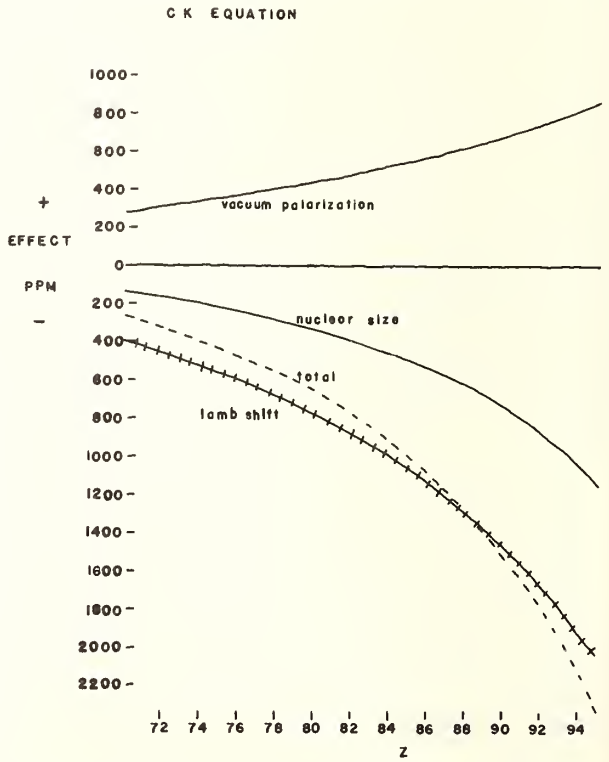


FIGURE 5. *The effect of various factors on the* L_{II}-L_{III} *splitting.*

is here considered to have the linear dependence on z shown by the solid line drawn through the points with $z<70$. The decrease in B_2 for $z>70$ is due to the nuclear size and quantum electrodynamic effects. The error bars on the figure represent the uncertainty in the value of B_2 for the given z due to the uncertainty in the corresponding value of the L_{II}–L_{III} binding energy difference. Figure 4 contains the information necessary to obtain an estimate for the magnitude of the Lamb shift in those elements with $z>70$. The difference between the calculated and experimental L_{II}–L_{III} energy level splitting is apportioned among the known causes such as nuclear size effect and vacuum polarization and the remainder is ascribed to the Lamb shift. Figure 5 illustrates this process and shows how much the Lamb shift affects the L_{II}–L_{III} binding energy difference.

While the process, sketchily reported here, is typical of the use of binding energy values in the study of quantum electrodynamic effects in heavy elements, it has some serious drawbacks. First, the analysis is dependent on an extrapolation (in the case of B_2), always a risky business. Second, it attributes the entire unexplained splitting difference to one cause, the Lamb shift. Furthermore, the resulting estimate of the Lamb shift is highly dependent on the quantity (in this case B_2) which is extrapolated. Nevertheless, it represents the best that this approach can do. Here is one case where the analysis would be greatly improved if one could increase significantly the precision with which the binding energy differences are known.

4. References

[1] Bearden, J. A., and Burr, A. F., Rev. Mod. Phys. **39**, 125 (1967).
[2] Bearden, J. A., and Burr, A. F., Atomic Energy Levels (U.S. Atomic Energy Commission, NYO 2453-1, Oak Ridge, Tenn., 1965).
[3] Burr, A. F., Bull. Am. Phys. Soc. **10**, 1221 (1965).
[4] Burr, A. F., and Liefeld, R. J., Bull. Am. Phys. Soc. **12**, 562 (1967).
[5] Burr, A. F., and Chamberlain, M. B., Bull. Am. Phys. Soc. **15**, 560 (1970).
[6] Richtmyer, F. K., Barnes, S. W., and Ramberg, E., Phys. Rev. **46**, 843 (1934).
[7] Burr, A. F., Advances in X-Ray Analysis **13**, 426 (Plenum Press 1970).
[8] Dev, B., and Brinkman, H., Bull. Am. Phys. Soc. **14**, 949 (1969).
[9] Johansson, P., Arkiv Fysik **18**, 289 (1960).
[10] Burr, A. F., and Liefeld, R. J., Bull. Am. Phys. Soc. **13**, 576 (1968).
[11] Chamberlain, M. B., thesis, New Mexico State University, 1970 (unpublished).
[12] Krishnan, T. V., and Nigam, A. N., Proc. Ind. Acad. Sciences **60**, 75 (1964).
[13] Burr, A. F., Bull. Am. Phys. Soc. **10**, 548 (1965).
[14] Christy, R. F., and Keller, J. M., Phys. Rev. **61**, 147 (1942).
[15] Shacklett, R. L., Rev. Mod. Phys. **30**, 521 (1958).

DISCUSSION

G. W. Series: As you pointed out or as is clear from your slide, before one can estimate the Lamb shift from any information of the sort you were discussing, one needs to know something about the nuclear volume correction. Now, this has been discussed before. You said it was well known. I wonder if you have a reference in your text or whether you can give us a reference now. I'm sure it is a most important feature of this calculation.

A. F. Burr: Yes. Shacklett has discussed this in a Review of Modern Physics article. The exact reference escapes me just now, but I have it in my briefcase, and I'll be glad to give it to you. (*Reference 15 in printed text.*)

G. W. Series: Thank you very much.

X-Ray/Optical Interferometry and X-Ray Fundamental Constants Measurements

Richard D. Deslattes

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

High quality synthetic crystal production has become well established during the past decade. Concomitant progress in defect crystallography allows for respectable attempts at characterization of residual defects. These developments suggest that crystals may again make significant contributions to our knowledge of the connection between microscopic and macroscopic quantities as for instance in the case of mass. Already available perfection in crystals has permitted the construction of large scale x-ray Moiré interferometers, static forms of which were first demonstrated by Bonse and Hart in 1965. It is evident that were similar devices arranged to provide for relative motion of their parts, and coupled to appropriate optical interferometers, crystal lattice parameters could thereby be measured in a fashion not depending on x-ray wavelength values. Reports presented at this conference will show several realizations of coupled x-ray and optical interferometers and certain initial results. While such first results are expected in the parts per million region, it is clear that one or two more orders are available with simple extension of present technology. This potentiality is in evident contrast with the situation of the "x-ray wavelength scale" whose internal refinement beyond the part per million level appears foreclosed by the obstinate tendency of x-ray lines to exhibit widths exceeding 300 parts per million. It is the aim of this brief review to examine some of the alternative responses which this situation might elicit regarding x-ray type measurements of the fundamental constants.

Key words: Atomic constants; lattice parameters; x ray.

1. Introduction

This report examines the status of x-ray measurements implicating numerical values for the constants in the light of recent developments in x-ray/optical interferometry [1]. Hopefully, this introduction can serve two ends: On the one hand it appears useful to try to examine what is going on in the x-ray business before this, a wider audience than the small cadre of practitioners. On the other hand, it seems also useful to try, even though the attempt may well err, to establish within the field a framework or ideology toward which we may attempt to orient our present and future efforts.

By way of an introductory reservation, I should acknowledge the difficulties which Prof. Bearden has discovered in the serious application of Bragg's law to real situations [2]. It would be a folly to deny the many problems which exist in reckoning a $\lambda/2d$ ratio from crystal goniometry.

There are, of course, those problems which follow from the residue of imperfection which persists in many, though not all, available diffraction crystals. Even granting intrinsic perfection, surface conditions have important effects. Although symmetric Laue diffraction angles are immune from index of refraction corrections, realization of this immunity requires stringent preparative procedures. Bragg geometry, to which one must ultimately turn at long wavelengths requires large corrections for index of refraction. This latter is most certainly not that following from the simple Snell's law estimates with an assigned average electron density. In the absence of evidence to the contrary, I take it that a conscientious application of the dynamical theory to wavelength measurements using well prepared "perfect" crystals can be satisfactory at least at the one part per million (ppm) level.

When a suitably articulated x-ray interferometer is joined with a sensitive form of optical interferometer it becomes possible to measure crystal lattice spacings in terms of visible wavelengths. Visible wavelengths are, of course, readily connected either with the present ^{86}Kr standard or directly with those lines from which the Rydberg is reckoned [3]. I feel that the potential contribution from this area is such that, at the very least, the corpus of x-ray measurements would regain significance even were they carried out with their traditional reference to an x-ray wavelength scale.

There is, however, the more exciting possibility of carrying out the classical measurements without reference to the present kind of x-ray wavelength scale and hence in a fashion not limited by the imprecision of this scale.

Realization of this possibility requires development of measurement algorithms not presently in the literature. To a certain extent, proposals for and development of these algorithms will influence those measurements which will be made. It thus appears

reasonable to attempt to set out at least one model for an ideological framework for such advanced measurements.

To this end, I shall first make a very brief resumé of the present status of the classical x-ray measurements with attention to the distribution of weight in their error budgets. Next, I shall describe, schematically, a hypothetical x-ray/optical interferometer measurement and indicate broadly the accuracy levels which may be realizable at present and in the nearby future. It becomes then of interest to consider how to integrate these possibilities into the scheme by which knowledge of the constants is ascertained.

At the most elementary level, we can attempt to see how such a measurement may be put to immediate use without requiring auxilliary measurements. Beyond this elementary and limited starting point it is possible to forsee several advancing stages of sophistication which lead ultimately to potentially highly accurate measurement algorithms requiring no external reference scale.

2. Structure of X-Ray Constants Measurements at Present

Measurements of primary interest are: (a) the high frequency limit of electron bremsstrahlung (h/e); (b) density and lattice parameter of a crystal (N_0); and (c) the electron-positron annihilation wavelength (h/mc). Since these measurements have, in fact, been carried out on an x-ray wavelength scale, their descriptions need to be augmented by inclusion of a suitable power of Λ, a multiplicative conversion factor from the x-ray scale to the conventional scale of length. (In fact, there is some confusion between the scales of extension and of wavelength; this, however, occurs at a level of refinement below that presently germane to the x-ray situation.) The conversion factor, Λ, thus becomes an equally necessary object of measurement (call it "d" in the above sequence) even though it has little theoretical interest.

There is another aspect of the wavelength scale business which has not been much emphasized, largely because, I suspect, it has not yet been a limiting factor in measurements. X-ray lines (at least the sharper ones) share a common full width at half maximum of about 300 ppm (parts per million). These lines do not have a known analytic profile and are observed under shot noise limited conditions [4]. These facts limit x-ray measurements, as presently performed, to the one ppm level or perhaps slightly beyond that.

It has therefore seemed desirable to examine the several x-ray/constants measurements in these terms: Are the measurements stopped at the part per million level for reasons other than the imprecision of the wavelength scale? Can the part per million imprecision of the wavelength scale be circumvented in each case by an alternative measurement algorithm? What are the steps which can be taken now and in the proximate future? This first part of my discussion is formulated with reference to an x-ray wavelength scale even though subsequent analysis may show that such reference is avoidable.

2.1. h/e From the High Frequency Limit

The shape of the high frequency limit (h.f.l.) from solids is not presently understood. That from gases is calculable. If solid state type peculiarities are assumed to be energy independent for sufficiently high energies, then differential type measurements are appropriate [5]. For the case of monatomic gas targets, there are no particular theoretical obstacles of principle precluding a shape calculation for a specified electron energy.

The most recent such measurement [6] used a mercury gas target and an x-ray wavelength reference. Principal limitations were, (estimates of probable error)

(1) recognition of the wavelength reference feature—10 ppm
(2) wavelength conversion factor—24 ppm
(3) electron energy scale—13 ppm
(4) as maintained/absolute electrical units—5 ppm
(5) shape of h.f.l.—12 ppm.

The last named problem might be somewhat easier in a rare-gas target, or two or more such gases might be used to test available theoretical models for the h.f.l. shape. (Naturally, quite large pumping capacity is required which is much more of a problem with rare gases than it was with mercury.) The electrical standards problem [4] does not differ from that encountered in the Josephson effect measurements [1]. It is a matter of taste whether kilovolts or millivolts are harder to measure to 1 ppm; in principle it would seem that kilovolts are likely easier.

The origin of the electron energy scale was conveniently established by noting the apparent ionization potential of Hg in the experimental chamber. Other excitation/ionization channel thresholds determined from optical data are available [7].

I hope to be able to show that items (1) and (2) can be substantially reduced by the results of x-ray/optical interferometry. This will apparently be effective to the extent of replacing the combined uncertainties of (1) and (2) by an uncertainty at or near one ppm. Thus the dominant errors would be rendered secondary while the remaining ones seem reducible to of the order of the electrical unit problem, e.g., less than 5 ppm.

2.2. $N_0\Lambda^3$ From Crystal Density and Lattice Parameter

The proposition underlying x-ray measurements regarding Avogadro's constant is that well ordered crystallization proceeds by symmetric replication of a fundamental (asymmetric) pattern unit to produce a macroscopic specimen whose extensive properties are connected with those of the molecular unit by

enumeration. Of the many questions which may be raised in regard to such a measurement, some address the crystal chemistry and characterization which are required to assure that symmetric replication is a believable operation at the one ppm level. Other questions may be raised even assuming problems of crystal perfection have been satisfactorily addressed. Structural and chemical perfection is quite a long story—too long, I feel for this presentation. A review of these aspects has been given [8] which includes model calculations. These suggest that present technology can, in fact, approach the one ppm level for symmetric replication. The remaining considerations include: isotopic abundance; density; lattice parameter and the scale on which it is reckoned. My discussion will emphasize these questions.

A recent measurement of this class by Ivars Henins and J. A. Bearden [9] was, in fact, oriented to use an external value of N_0 together with measured values of densities and lattice parameters to infer a value for Λ. For consistency, I choose to examine these data as though the object was to measure N_0, in which case, the error budget appears as follows:

(1) isotopic abundance variability—10 ppm
(2) density of water—3 ppm
(3) crystal chemistry (estimate from reference [8])—2 ppm
(4) x-ray wavelength—2 ppm
(5) conversion factor between x-ray and optical wavelength—20 ppm.

Clearly, the problem of isotopic abundance shares dominance with the wavelength problem. If silicon persists as a candidate crystal for such measurements then a substantial degree of isotopic enrichment is required before current mass spectrometric technique [10] can yield an effective atomic weight good to 1 ppm. Other crystals not suffering from the isotope problem as badly as Si are available in qualities approaching that required for this measurement [8]. It is a matter of expediency which of these routes is preferable at a particular time.

If the isotopic abundance problem is dealt with in one way or the other, there then remains principally the problem of a metric lattice parameter. This will be treated in discussing the x-ray/optical interferometer. Additionally there are questions of how to make the connection between macroscopic and microscopic values. For example, the Bristol/NPL approach is to measure the volume of a crystal and weigh it. At NBS, Bowman [11] is using noncrystalline density standards and coupling to crystal densities via hydrostatic weighing. One of these alternatives has yet, in my opinion, to emerge as clearly superior to the other.

2.3. $h/m_e c$ From Electron-Positron Annihilation

The wavelength from electron-positron annihilation has been a useful approach to the electron's Compton wavelength. A principal problem has been large spectral width conveyed to the annihilation photons by the initial center of mass momentum of the e^+, e^- pair. An approach to reducing this problem will be discussed by Sauder at this conference [12].

The work of Knowles [13] will serve to illustrate the sort of error distribution encountered in this type of measurement.

(1) line location (including crystals)—30 ppm
(2) angle measurement—10 ppm
(3) x-ray wavelength—2 ppm
(4) conversion factor Λ—20 ppm.

It is evident that the crystal diffraction width problem is considerably reduced in the results reported by Assche et al., in their bent crystal results reported at this conference [14]. It seems that this problem ean be even further reduced in a double flat crystal experiment [12]. Basically, the experiments of Knowles, and of Assche et al., compare annihilation wavelength with one or more x-ray standards. The imprecision of these ultimately limit the measurements while they are proximately limited by the conversion factor.

3. X-Ray/Optical Interferometry

Although there are other possibilities, e.g., Michelson analogs, etc., work reported at this conference has employed nondispersive, Moiré type x-ray interferometers [15, 16, 17]. These have a great many technical advantages not the least of which is that the signal accompanying motion depends only on the interplanar separation in the crystal.

A hypothetical joint x-ray/optical arrangement is illustrated schematically in figure 1. In operation, relative motion, not necessarily continuous, is given one part of the x-ray interferometer together with its attached optical element with respect to the other pair of joined elements. It is obvious that there are formidable technical obstacles buried in this simple verbal description. Some of the technical develop-

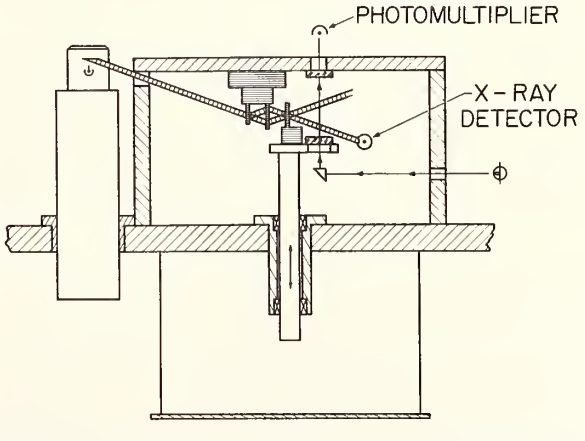


FIGURE 1. *Hypothetical arrangement for a coordinated x-ray/optical interferometer.*

ments required to effect such an operation will be described in the following papers [15, 16, 17].

For purposes of this discussion, I take it that the result is at hand. Presently expected results are of the form "in this particular Si specimen, 1648.xxx," (220) repeat distances occurred for each passage of one-half optical wavelength of 6328 Å light from a Lamb dip stabilized $^{3}He^{22}Ne$ laser." Results available or shortly expected suggest one ppm accuracy for this statement is readily achievable. With some further work, results may be anticipated having two orders of magnitude greater accuracy.

The question to which I now turn is: granting available results of this sort, what should we do next? How can these evidently important data be integrated into the structure from which knowledge of the constants is derived?

3.1. Initial Approaches

If it is assumed that all silicon crystals are alike to a significant accuracy, then the above result can be combined with those reported by Henins and Bearden [9] to obtain both Λ and N_0. The extent to which such a procedure may be in error is suggested by the results obtained from two groups of Si crystals of rather excellent quality [18]. These were obtained from two suppliers. Samples from each specimens exhibited density values consistent to within 2 ppm while the groups differed by more than 10 ppm. Indeed, differences in the 20–30 ppm range had been found earlier [18]. This appears sufficient to foreclose the "generic silicon" approach. It likewise suggests the inadvisability of attempting to define a new x-ray wavelength scale based on "the lattice parameter of perfect silicon at 25 °C."

At the very least then, one must attempt to measure one or more x-ray reference wavelengths either with a part of the interferometer or with some other crystal(s) which have been related to it in a fashion which does not degrade the accuracy derived from the interferometric measurement. If such a procedure were carried out, the results would be quite useful at present but would be limited by the imprecision inherent in the x-ray seale insofar as it is embodied in one or several x-ray lines.

To the extent that our interest is focussed on constants measurement, it is pertinent to note that *no x-ray wavelength enters any of the fundamental constants measurement in an essential way.* We are thus led to consider alternative measurement chains, not influenced by the lack of sharpness of the x-ray lines.

3.2. Distribution of an Interferometer Result

The physical form of an interferometer crystal is not particularly apt for density measurements or for that matter either high energy (h/mc) or low energy (h/e) spectroscopy. We need measurement schemes in which the interferometer calibration can be transferred to more appropriate samples for subsequent measurements. Any useful scheme must work to accuracy levels below one ppm, that is, it must depend at most weakly on x-ray wavelengths, if they are used at all. Interesting schemes should also be immune to first order corrections for index of refraction since these are quite trying [3].

There appear to be at least three schemes which satisfy these desiderata. Each appears to have certain features which are quite attractive for one purpose but somewhat disfunctional for at least one of the others. Pursuit of two independent chains to one end might be reassuring with respect to systematic error. The three procedures whose descriptions follow have all appeared in the literature. However, they have not, to my knowledge, been compared with one another in the context of this diseussion.

a. Quasi Nondispersive Lattice Parameter Transfer

Hart has proposed [19] a dual-source transfer measurement for lattice parameters which is especially useful in the case of near equality. Figure 2 is intended to suggest a situation in which crystals C_1, C_2, and X have closely the same lattice parameter. Moreover it is assumed that C_1 and C_2 are parallel by virtue of adjustment or because they are carved from a single block. These "first crystals" C_1 and C_2 may be used in either surface reflection or in transmission. A nominally "equal" crystal introduced at X produces maxima in the detectors at only slightly different angular positions. For simplicity, we can assume that C_1 and C_2 are closely parallel so that the difference in angular coordinate of the two maxima depends only on the grating space difference. The maxima will then occur at angles

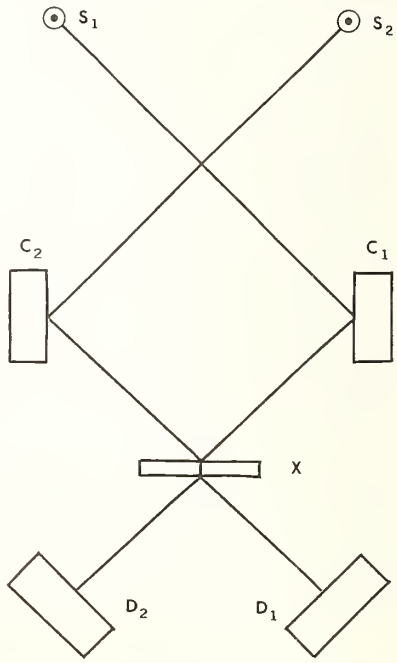


FIGURE 2. *A lattice parameter transfer measurement suitable for near equality of lattice parameters of the crystals* C *and* X. Standard and "unknown" crystals are exchanged in position X.

differing by $\delta\theta$ such that $d_x = d_c \pm \lambda \cot\theta\delta\theta$. Thus, for near equality of the d's, a poor knowledge of λ suffices. Likewise, the angular accuracy required is far less than in a direct method. The measurement protocol would interchange interferometer and sample crystals in the "X" position. To the extent that both measurements can be made in transmission, there is no refractive index correction.

For measurements at long wavelengths (say $\lambda > 1$ Å), crystals are generally required to work in surface reflection (Bragg case). Even here, there is an elegant procedure for evading the index of refraction correction. The transmission measurement does not depend on wavelength so one may choose to use the λ which is contemplated for subsequent surface reflection. The transmission "standard" (the calibrated part of an interferometer) is then replaced by a reflection "unknown." The differing $\delta\theta$'s then relate *the* grating space of the "standard" to the *effective* grating space of the "unknown" (at that λ and in that order of diffraction).

b. Use of a Monolithic Monochromator

One can easily produce a beam of x-rays more monochromatic than natural x-ray lines by selecting a part of such a line in a double crystal monochromator. Although use of a conventional instrument for such purpose is possible, it would be expensive and require a rather rigorous measurement protocol to insure absence of instrumental drift. On the other hand, a double crystal monochromator carved from a single crystal of decent material is manifestly stable (the interplanar or "intercrystal" angle depends only on symmetry for a cubic crystal and for certain plane pairs in other systems possessing at least one axis of symmetry).

Such a scheme is practical only if the "eigen-lambda" of some such device lies within a prominent characteristic x-ray line, since only in this case can an intense beam be made available for the calibration measurements. To see whether this kind of scheme might go, some 10,000 eigen-lambdas were calculated for Ge and Si plane pairs having $h^2, k^2, l^2, < 100$ [20]. This list was then compared with the more prominent lines in the spectra of those elements which are tractable as anode materials [21]. About 20 interesting coincidences occurred, one of which is illustrated in figure 3 for Ge. The error between the eigen-lambda and the peak of Co $K\alpha_1$ is 60 ppm which can be reduced, if necessary, by thermal tuning.

The protocol envisioned would use the calibrated interferometer segment to determine the wavelength being prepared by the monolithic monochromator. Thereafter, other crystal specimens could be inserted in the beam and angles measured between their diffraction maxima to establish their (effective) grating spacing. As compared with the previous approach, this one evades the dual source requirement at the price of having to measure angles as accurately as we wish the result to be meaningful. This inelegance can be overcome by hard work.

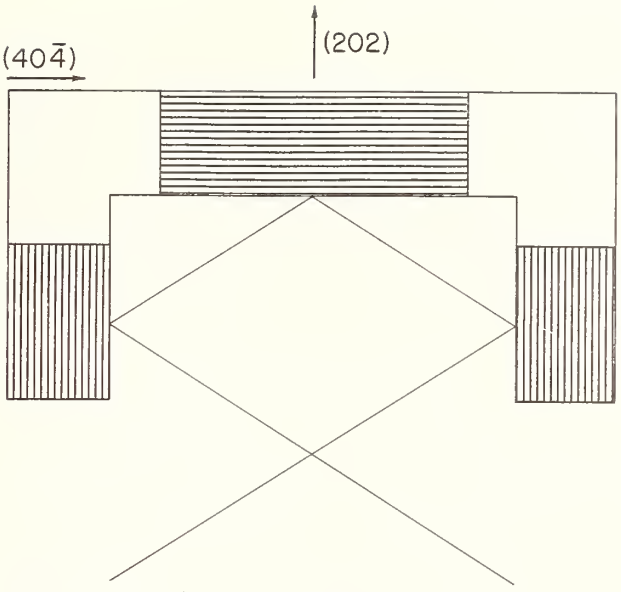


FIGURE 3. *Diagram of* Ge *monochromater whose pass band lies within* CoKα_1, *see text.*

There is a dividend to this approach which follows from the fact that, in order for the monolith to be useful in calibrating crystals, its eigen-wavelength must be known. This information is a by-product of the steps necessary for lattice parameter transfer. It may then be advantageously introduced into an h/e type measurement.

c. Nuclear Ray Sources

Low energy nuclear γ rays are quite sharp [22], in fact, one can generally find lines two orders of magnitude sharper than x-ray lines. To use such sources as effective substitutes for x-ray tubes one need only be willing to deal with multikilocurie radioactive sources. This will usually involve reactor irradiation for source production. It will be necessary to provide a fairly rigorous protocol for safe use of such sources.

Application of such sources to a calibration chain beginning with an interferometrically calibrated specimen requires again accurate angular measurement. A desirable byproduct of this exercise is that at least one nuclear γ radiation comes to have its wavelength tied to visible standards. From only this base, and with no further accurate large angle measurements, it is possible to connect a significant group of nuclear emission wavelengths (ideally a group of near harmonics) with the one which is tied to visible light.

This approach has much to recommend it since the γ rays are manifestly independently reproducible. An harmonic chain to 0.5 MeV is easy to find. If the source handling problem could be reduced to routine, sharp secondary standards would be freely available. (In connection with the hot source problem, it should be realized that x-ray tubes are equally active, their only claim to privilege is that they can be turned off.)

4. Concluding Speculations

I hope to have been able to suggest some of the potential for x-ray/optical interferometry in bridging the three order of magnitude gap between x-ray and optical wavelengths. No development in the past 40 years has been quite as hopeful regarding this connection problem. I hope also to have opened a discussion or a debate on how best to exploit this new technology to improve our basic knowledge.

5. References

[1] All reviews to date exclude x-ray interferometry for obvious reasons; the most recent is Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[2] Bearden, J. A., these Proceedings.
[3] X-ray interferometry begins with Bonse, U., and Hart, M., Appl. Phys. Letters **7**, 99 (1965). Static reassembly of separated parts was demonstrated by Bonse, U., and te Kaat, E., Z. Physik **214**, 16 (1968), and elastic scanning of an articulated monolith by Hart, M., J. Phys. 1D, Scr. 2, 1405 (1968). A mixed x-ray/optical assembly of suitably refined character for these measurements was reported in late 1969: Deslattes, R. D., Appl. Phys. Letters **15**, 386 (1969).
[4] Thomsen, J. S., and Yap, F. Y., J. Res. Nat. Bur. Stand. (U.S.), 72A (Phys. and Chem.), No. 2, 187–205 (1968).
[5] Bearden, J. A., private communication.
[6] Spijkerman, J. J., and Bearden, J. A., Phys. Rev. **134**, A871 (1964).
[7] See for example, Simpson, J. A., Menendez, M. G., and Mielczarek, S. R., Phys. Rev. **150**, 76 (1966).
[8] Deslattes, R. D., and Peiser, H. S., and Bearden, J. A., and Thomsen, J. S., Metrologia **2** (1966), p. 103.
[9] Henins, I., and Bearden, J. A., Phys. Rev. **135**, A890 (1964).
[10] I am indebted to W. R. Shields for discussion on this subject.
[11] Bowman, H. A., private communication.
[12] Sauder, W. C., these Proceedings.
[13] This work is reviewed by Knowles, J. W., in "Crystal Diffraction Spectroscopy of Nuclear γ-Rays." Alpha-, Beta- and Gamma-Ray Spectroscopy, Chapter IV, Vol. 1 (1965) (North Holland, Amsterdam).
[14] Van Assche, P. H. M., Van den Cruyce, J. M., Vandenput, G., Baader, H., Koch, H. R., Breitig, D., and Schult, O. W. B., these Proceedings.
[15] Curtis, I., Hart, M., Milne, A. D., and Morgan, I., these Proceedings.
[16] Bonse, U., and te Kaat, E., these Proceedings.
[17] Deslattes, R. D., these Proceedings.
[18] Bowman, H. A., and Schoonover, R. M., with Appendix by Jones, M. W., J. Res. Nat. Bur. Stand. (U.S.), 71C (Eng. and Instr.), No. 3, 179–198 (1967).
[19] Hart, M., Proc. Roy. Soc. **A309**, 281 (1969).
[20] Deslattes, R. D., Appl. Phys. Letters **12**, 133 (1968).
[21] Bearden, J. A., Rev. Mod. Phys. **39**, No. 1, 78 (1967).
[22] Greenwood, R. C., Helmer, R. G., and Gehrke, R. J., Nucl. Instr. and Meth. **77** (1970), pp. 141–158.

DISCUSSION

P. H. M. Van Assche: The problem to know which transitions have to be measured absolutely with such a device as you are developing depends on where the needs are greatest. When we reconsider the problems Professor Wapstra and, of course, many other people encounter for having absolute gamma-ray wavelengths, I think this would be at least a very useful approach to try and get absolute wavelengths for gamma rays.

R. D. Deslattes: I fully agree. As a matter of fact, we started this thing from a rather narrow point of view. In subsequent years it has transpired that the market for decent gamma ray measurements is quite large indeed, and I hope that we can contribute to this. For example, there is a particularly large and significant market for references for mesic x-rays where one ultimately gets at the pion/nucleon interaction and in first order checks electrodynamics.

M. Hart: I'd like to make a comment on crystal perfection because I think the paper has been a little bit misleading there. Dr. Deslattes stated that they have measured density, and the density measurements comparing float-grown crystals of silicon with Czochralski-grown silicon crystals differ in lattice parameter by 10 parts per million, even though both were homogeneous in themselves. Is that the figure?

R. D. Deslattes: No, you're off by a factor of three. I quoted the *densities* as being different by 10 parts per million, the lattice parameters by 3 parts per million.

M. Hart: That's fine. The immediate reaction from a few people around me who happen to be interested in defects was, "That's not surprising." And, in fact, it's well known that Czochralski silicon contains oxygen, and measurements have been made of correlations between oxygen and lattice parameter.

R. D. Deslattes: I agree with that remark, but not—I can't use the trade name—but not in that stuff that comes from down there. (*Laughter*)

M. Hart: Okay.

Another piece of information which I think is rather more realistic in state of the art: This is a publication by Buschert and coworkers [J. Appl. Phys. 39, 4365(1968)] from Dow-Corning Corporation. They grew silicon which as far as we know and as far as they could analyze is the best ever grown. On six different occasions they grew six different crystals, compared their lattice parameter, and the total spread was plus or minus 2 parts in 10^8. I think if people start thinking of that sort of material they may well have a rather different view of what might be done in the way of distributable standards.

R. D. Deslattes: I fully agree about distributable standards, Mike, but the problem is isotope variability. I don't think these are defect-conditioned things. I think that they are geologically conditioned things. If it's economically possible to turn the group you just described (whose name I won't mention) loose on this problem with separated isotope silicon, then I withdraw my objection. Okay?

M. Hart: Okay.

R. D. Deslattes: But the manifest independent reproducibility of a nuclear gamma ray does have a certain esthetic appeal.

M. Hart: Oh, that's fine. (*Laughter and applause*)

Measurement of the Compton Wavelength of the Electron

P. H. M. Van Assche

Studiecentrum voor Kernenergie, S.C.K./C.E.N., B-2400, Mol, Belgium

J. M. Van den Cruyce and G. Vandenput

Katholieke Universiteit te Leuven, B-3030 Heverlee, Belgium

and

H. A. Baader, H. R. Koch, D. Breitig, and O. W. B. Schult

Physics Department of the Technical University, Munich, Germany, and Research Establishment, Risö, Denmark

Using a bent-crystal γ-diffraction spectrometer, the Compton wavelength of the electron has been measured relative to the W $K\alpha_1$ wavelength (expressed in X-units). The annihilation line occurred in presence of the neutron-capture γ-ray spectra of Ta. The good resolution of the spectrometer allowed a discrimination against some capture γ lines superposed upon the annihilation line. The linearity of the spectrometer has been checked with intense capture γ transitions, observed in different diffraction orders. With this technique a final precision of better than 30 ppm has been obtained. The final value is $\lambda_c = 24.212\ 69 \pm 0.000\ 77$ XU, relative to $\lambda = 208.576\ 7 \pm 0.000\ 3$ XU for the wavelength of the $K\alpha_1$ x-ray of ^{182}W.

Key words: Compton wavelength of the electron; γ-diffraction spectrometer; electron annihilation; X-unit.

1. Introduction

In order to measure the Compton wavelength of the electron, the 511 keV annihilation peak has been compared to the W $K\alpha_1$ line with the bent crystal diffraction spectrometer from the Technical University, Munich. This spectrometer is located at the DR3 reactor of the Danish Research Establishment. Both annihilation and x-ray lines are present in the γ-spectrum following the neutron capture in Ta.

2. Experimental Method

The gamma radiation is measured with a DuMond-type bent crystal spectrometer. The crystal, a 4 mm thick quartz lamina cut perpendicularly to the (502) and (110) planes, is bent to a radius of 5.8 m by cylindrical clamping blocks. The gamma radiation is reflected from the (110) planes with grating constant $d \approx 2.5$ Å. This spectrometer has recently been made automatic [1]. The resolution of this spectrometer is such that some (n, γ) lines, superposed upon the annihilation peak, can be observed. As many γ-transitions are observed in different diffraction orders between 50 to 1000 keV, the linearity of the spectrometer can be checked by comparing the actual diffraction angles in different orders with their exact values determined by the Bragg law $n\lambda = 2d \sin\theta$. With this method the non-linearity of the spectrometer can be corrected to better than 0.17″ over the energy region in which we are interested.

3. Results

3.1. Definition of the X-Unit

According to a proposal of J. A. Bearden [2] the wavelength of the W $K\alpha_1$ x-ray can be accepted as a basis for the definition of x-ray wavelengths. At that moment, Bearden proposed to put the wavelength of this W $K\alpha_1$ transition equal to 0.209 010 0 Å* (angström star) reposing on a recommended value [3] for the conversion constant $\Lambda = 1.002\ 076 \pm 5$ ppm, being the ratio of a wavelength expressed in angström to the same wavelength expressed in kilo-X-units. With this numerical value of Λ, he expected the new unit Å* to be equal to 1 Å (10^{-10} m), within ± 5 ppm.

Since another conversion factor appears then [4], we prefer to use the values proposed by Bearden, with the only difference that the wavelength of the W $K\alpha_1$ line is expressed in X-units; this yields $\lambda(W\ K\alpha_1) = 208.5770$ XU.

This convention will allow the authors to express their results in a convenient way. In our spectrum the x-rays from W are produced after internal conversion in the ^{182}W nucleus. A small isotope shift

$$\Delta\lambda = 0.28 \pm 0.04 \text{ m XU}$$

has been observed between $K\alpha_1$ x-rays of ^{182}W and ^{184}W (see [5]). This shift should normally not be taken into account as it is about ten times smaller than the standard errors on wavelengths in this region. In this case however, the wavelength, being a defined value, is used without error. On the other hand, the isotopic difference between the natural composition of W (atomic mass 183.85) and the ^{182}W isotope, is nearly equal to the isotopic difference between the 182 and 184 isotopes as used by Chester and Boehm. Therefore we will propose to use a value of

$$\lambda = 208.5767 \pm 0.000\,3 \text{ XU}$$

for the wavelength of W $K\alpha_1$ when it is produced in the decay of ^{182}Ta to ^{182}W.

3.2. Capture γ-Lines Superposed on the Annihilation Line

The main problem in determining the annihilation line in this Ta(n, γ) experiment is due to the fact that some (n, γ) transitions are superposed on its shape; the strongest of them is a 509.96 keV-line with nearly 7 percent of the intensity of the annihilation line in this experiment. A typical shape is given in figure 1. All superposed lines are observed in different diffraction orders; therefore their existence is real and they may be subtracted from the annihilation line.

3.3. Wavelength Determination

The wavelengths of the annihilation lines observed in two separate spectra are given in table 1.

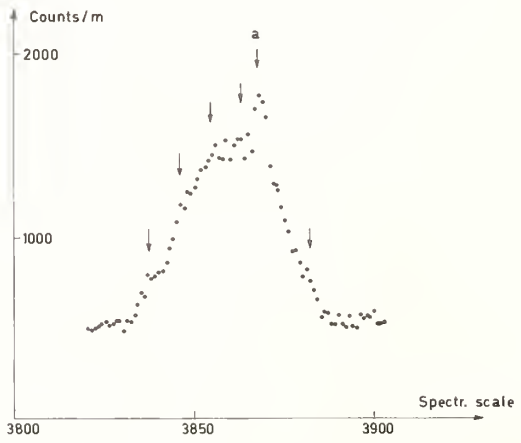


FIGURE 1. *Annihilation line with* (n, γ) *transitions superposed upon it; the* (n, γ) *transitions which are observed in different diffraction orders are indicated with an arrow.*
The 509.96 keV (n, γ) line is indicated by arrow a.

TABLE 1. *Wavelengths of annihilation peaks, observed in* Ta(n, γ) *spectra*

	Wavelength	Standard deviation
Diffraction order	λ (XU)	σ (m XU)
1	24.219 51	6.00
2	24.211 74	3.35
3	24.210 03	2.25
4	24.211 59	3.00
5	24.214 06	1.54
1	24.211 65	7.00
2	24.213 50	3.20
3	24.213 50	2.40
5	24.211 90	1.50

Due to an important delay in the analysis of a third Ta(n, γ) spectrum, these data could not be included in our present results.

The mean value obtained from table 1 is $\lambda =$ 24.212 61 XU. The standard deviation on this mean value was calculated by internal and external consistency formulae, yielding respectively $\sigma_{int} = 0.77$ m XU and $\sigma_{ext} = 0.54$ m XU; it is normal to choose the largest value. The non-linearity is actually the limiting factor of our experiment.

The kinetic energy of the center of mass of the position-electron pair can be calculated from the line broadening of the annihilation peak: it is about 3.2 eV. The electron rest mass energy differs from the observed annihilation quanta by half of this energy; therefore the observed wavelength has to be increased by about 3 ppm. Finally we obtain for the Compton wavelength of the electron

$$\lambda_c = 24.212\,69 \pm 0.000\,77 \text{ XU}.$$

4. Discussion

This method of measuring the Compton wavelength of the electron in a Ta(n, γ) spectrum has first been applied by Knowles [6], who used a double flat crystal diffraction spectrometer. The main difference with Knowles' experiment is that the resolution of the bent crystal spectrometer at Risö is such that the disturbing (n, γ) lines can be observed and subtracted accordingly. These γ-transitions, especially the intense line at 509.96 keV could be responsible for the difference between the two results of Knowles [6]. Indeed, the result obtained with the Ta(n, γ) spectrum is shifted to lower energies with respect to the other result obtained with a H(n, γ) spectrum.

Before comparing our result with those published previously the latter should be reduced to the same X-unit scale. The two results of Knowles [6] are quoted in [4] as

$\lambda_c = 24.212\,63\ (92)$ XU (H capture)

$\lambda_c = 24.214\,21\ (36)$ XU (Ta capture)

with respect to a $\lambda(W\ K\alpha_1) \equiv 208.5811$ XU. Taking the wavelength for $\lambda(W\ K\alpha_1)$ as in section 3.1. We obtain

$\lambda_c = 24.212\ 12\ (92)$ XU (H capture)

$\lambda_c = 24.213\ 70\ (36)$ XU (Ta capture).

As a conclusion it may be stated that our result strengthens that of Knowles taken with the $H(n, \gamma)$ spectrum.

5. Conclusion

The described method of comparing the annihilation wavelength to a standard x-ray wavelength in (n, γ) spectra seems to rely on sound principles, provided

—the linearity of the spectrometer can be checked carefully;

—the resolution allows subtraction of disturbing (n, γ) transitions upon the annihilation peak.

The linearity check—the main source of error in our experiment—will be improved in the near future, bringing the accuracy of λ_c in even more useful ranges.

6. Acknowledgments

This study has been performed under a joint program K.U.L.–S.C.K./C.E.N. We are indebted to the directions of both institutes for making possible this collaboration. The directors of the A.E.K., Risö, and the T.U., Munich are gratified for continuing interest in and sponsorship of the bent crystal measurement at Risö. The authors like to mention the kind hospitality they enjoyed during their stay at Risö.

7. References

[1] Koch, H. R., Baader, H. A., Breitig, D., Mühlbauer, K., Gruber, U., Maier, B. P. K., and Schult, O. W. B., in Proceedings of the International Symposium on Neutron Capture Gamma-Ray Spectroscopy, I.A.E.A. Vienna (1969), p. 65.

[2] Bearden, J. A., Phys. Rev. **137**, B455 (1965).

[3] Bearden, J. A., Phys. Rev. **137**, B181 (1965).

[4] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[5] Chester, R. B., and Boehm, F., Phys. Rev. **166**, 1206 (1968).

[6] Knowles, J. W., in Proceedings of the Second International Conference on Nuclidic Masses, W. Johnson, Ed. (Vienna, 1964), p. 113.

There are men at loose in this nation
Who apply Josephson with great concentration.
But what they didn't say
Earlier in the day
Is that alpha results from annihilation.

WILLIAM C. SAUDER

A Positron Annihilation Experiment for Determining the Fine Structure Constant

William C. Sauder

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

A measurement of the Compton wavelength to one part per million determines the value of α to one-half part per million. In our experiment, the Compton wavelength is determined by measuring the wavelength of the gamma radiation produced in two photon annihilation of positronium. A Bragg angle measurement to one ppm will be carried out with a two crystal Laue transmission spectrometer capable of angular measurements accurate to 10^{-3} arc s. The annihilation line will be narrowed by allowing the positronium decay to occur in low temperature helium gas in the presence of a magnetic field.

Key words: Angle measurement; Compton wavelength; double crystal spectrometer; fine structure constant; positron-electron annihilation.

1. Introduction

When two photon positron-electron annihilation occurs under idealized circumstances in which the two incoming particles are at rest (in the laboratory frame of reference), the total energy involved in the process is

$$E=2mc^2 \tag{1}$$

where m is the electron (or positron) mass. Since each photon carries off energy $E/2$, then the photon wavelength can be computed from

$$hc/\lambda=mc^2. \tag{2}$$

Solving eq (2) for λ, one finds that the annihilation wavelength is

$$\lambda=h/mc. \tag{3}$$

Thus the wavelength of the photons produced in the process is the Compton wavelength of the electron.

In light of the current situation in the state of the atomic constants, it is revealing to rewrite eq (3) in a different form as follows:

$$\lambda_c=\tfrac{1}{2}(e^2/\hbar c)^2(4\pi\hbar^3c/me^4)=\tfrac{1}{2}(\alpha^2/R_\infty) \tag{4}$$

where λ_c, α, and R_∞ are the Compton wavelength, the fine structure constant, and the Rydberg constant, respectively. Since the Rydberg constant is known to an accuracy far better than the presently attainable uncertainty in the measurement of λ_c, then the determination of λ_c from a positron annihilation experiment is equivalent to a measurement of the fine structure constant:

$$\alpha=(2R_\infty\lambda_c)^{1/2}. \tag{5}$$

In this paper I shall describe the progress in such an annihilation experiment. In particular, I shall (1) discuss the rather complex annihilation photon source that we will employ in order to reduce Doppler widths to a value consistent with the planned precision of the experiment and (2) report upon the sensitivity of the spectrometer that we have constructed to carry out the measurement. These two elements of the experiment—the narrow annihilation line source and the spectrometer of very high angular sensitivity—will permit a measurement of λ_c to an accuracy of one ppm (part per million), and thus a determination of the fine structure constant to a refinement of 0.5 ppm. Note that such a determination at the part per million level is free from any quantum electrodynamic corrections; thus the experiment provides a fine structure constant determination that is independent from and competitive with the other α-experiments carried out in recent years.

2. Problems Peculiar to This Precision Measurement

Several annihilation experiments [1, 2, 3] have been carried out in the past with the intent of

reducing measurement errors to a level such that the results would yield viable input data for atomic constants adjustments. If, however, one wishes to reduce uncertainties to the one ppm level, several problems arise that require significant changes in approach compared to the earlier work. The general method of the experiment remains the same, namely to measure λ_c by determining the Bragg angle at which the annihilation line is diffracted by a crystal spectrometer:

$$\lambda_c = 2d \sin\theta. \tag{6}$$

The development that makes possible a λ_c measurement of the accuracy contemplated is the possibility of employing as diffracting elements perfect (silicon or germanium) crystals for which the grating spacing, d, is found by comparing the crystals to a scanning x-ray interferometer [4] by means of a quasi nondispersive transfer experiment [5].

Since λ_c is approximately 0.024 Å, the Bragg angles θ for Ge and Si will be of the order of 1° or less. Thus a 1 ppm determination of λ_c requires the measurement of θ to a precision of 10^{-3} s of arc. This in turn establishes the angular noise levels allowable in the spectrometer, and furthermore requires that one employ optical interferometric techniques for the angle measurement. An instrument that satisfies these requirements has been completed; it will be described in section 4 of this paper.

Since we require not only an angular precision of 10^{-3} arc s, but also an absolute angular accuracy of this refinement, one must also provide for suitable calibration procedures for the angle measuring interferometer. We plan to use at least two redundant calibration schemes, one involving a direct measurement of the parameters of the interferometer, and the other based upon the diffraction of calibrated nuclear gamma line sources in our spectrometer as described by Deslattes [6].

If one requires a 10^{-3} arc s measurement of the location of the annihilation line, then an observed line width of at most 0.1 arc s is desirable. One of the limiting factors in previous λ_c measurements has been the fact that observed linewidths were more than a factor of ten greater than this. For perfect Si and Ge crystals, the diffraction widths at 0.5 meV are 0.1 s or less. Thus one now requires that the Doppler broadening of the annihilation line be reduced to this same order of magnitude.

If at the time of annihilation the positron-electron system carries a nonzero momentum, then conservation requires that the emerging photons carry off this momentum. As a result the annihilation line is Doppler broadened [7]. This rules out the observation of the annihilation of free positrons in this experiment, since this leads to linewidths of thousands of ppm wide. The only alternative is to form the hydrogen-like atom, positronium, and subsequently to thermalize the Ps atoms in a moderator sufficiently cold to reduce the Doppler widths to a suitable value. Our calculations show that a moderator temperature of 10 K will result in an annihilation line having a width of 50 ppm. The configuration of the source that we envision for the experiment is discussed in the next section.

3. The Source

The positron source for the experiment will be ^{64}Cu produced by neutron activation of the natural isotope ^{63}Cu in the NBS Reactor. To obtain a count rate at the detector of several hundred counts per minute at the annihilation peak will require a source activity of approximately 10 kilocuries. Because of the massive shielding requirements of the source and its relatively short half-life (12.88 h), the spectrometer will be located in a laboratory area adjacent to the reactor room.

As mentioned in the previous section, positrons from the ^{64}Cu source will be allowed to form positronium. The hyperfine splitting of the ground state of Ps amounts to approximately 8×10^{-4} eV, which is a negligible energy difference as far as this experiment is concerned. But it is of considerable importance that the upper state is triplet in character, has a mean life against annihilation of 1.387×10^{-7} s, and undergoes three photon annihilation. The lower state is singlet, has a mean life of 1.245×10^{-10} s, and decays into two photons. For the λ_c measurement, three photon decay is useless since this mode produces a continuous spectrum from 0 up to 0.511 meV.

Helium has been chosen as the thermalizing medium in our experiment both because of the low temperature requirements for the moderator and because Ps atoms are able to exist for their full natural lifetime in He (whereas in many other media, Ps lifetimes are shortened by processes such as "pick-off" or magnetic "spin flip"). Furthermore, we must employ He densities below 0.08 gm/cm^3, because, at higher densities, the exchange energy repulsion between Ps and He causes individual Ps atoms to be trapped in "bubbles" of radius 20 Å before reaching thermal equilibrium with the He. Thus the moderating medium is gaseous helium in the neighborhood of 5 K. We have estimated the time required to establish thermal equilibrium using a simplified classical hard sphere model [8]; the thermalization time for He gas under the prescribed conditions is 7×10^{-9} s.

It is immediately apparent that singlet Ps is not sufficiently long-lived to allow thermal equilibrium to be established with the moderator. This then requires that the source be immersed in a magnetic field. In the presence of a magnetic field, the states are mixed [7]; the wave function corresponding to the triplet $m=0$ state becomes, upon application of a magnetic field H,

$$\Psi_{10} = au_{10} + bu_{00} \tag{7}$$

where u_{10} and u_{00} are the $m=0$ wave functions for the zero field states, and the coefficients a and b are functions of H. The function Ψ_{10} is a normalized wave function; for $H=0$, $a=1$ and $b=0$. The mean

annihilation rate for the state Ψ_{10} is

$$\gamma_{10}=a^2\gamma_T+b^2\gamma_S \tag{8}$$

where γ_S and γ_T are, respectively, the singlet and triplet annihilation rates for the zero field states. The branching ratio, B_{10}, which is the ratio of the number of two photon to three photon decays out of the γ_{10} state is simply

$$B_{10}=b^2\gamma_S/a^2\gamma_T. \tag{9}$$

Dividing eq (8) by $a^2\gamma_T$ yields

$$(1/a^2)\,(\gamma_{10}/\gamma_T)=1+B_{10}. \tag{10}$$

Thus for a branching ratio of unity, the ratio of the mean life of the state Ψ_{10} to the mean life of the zero field triplet state can be written

$$\tau_{10}/\tau_T=1/2a^2. \tag{11}$$

Now since $\gamma_S/\gamma_T\approx 10^3$, then eq (9) indicates that B_{10} reaches the value unity when the ratio b/a is very small. Thus one can cause half of the Ps atoms in the triplet $m=0$ state to decay via the two photon channel with only a 50 percent loss in lifetime. Calculations show that the quenching field for this to occur is 2177 G [7], as shown in figure 1.

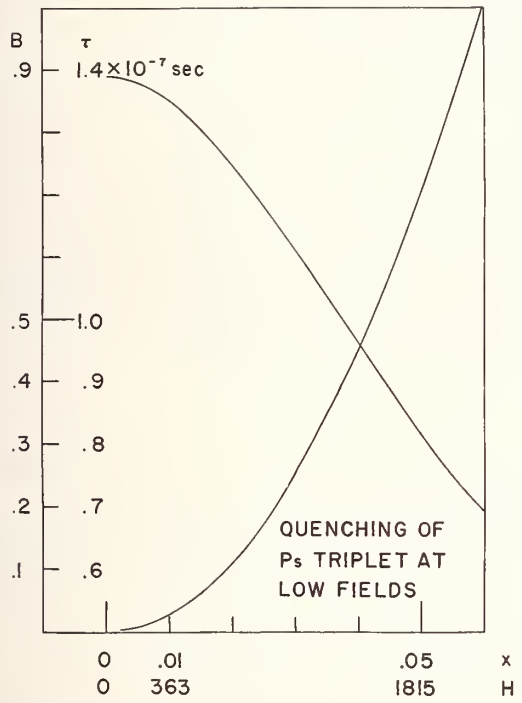


FIGURE 1. *Quenching of triplet positronium at low fields.* Falling curve is mean life of the state, the rising curve the branching ratio B_{10}.

In summary, the annihilation source will consist of a 10 kCi ^{64}Cu source immersed in He gas at 5 K in a magnetic field of 2 kG. The lifetime of the quenched state will be approximately 7×10^{-8} s; the thermalization time under these conditions is 7×10^{-9} s. As a result the annihilation line will be narrowed to approximately 30 ppm half width at half intensity.

4. The Spectrometer

The double crystal spectrometer that has been designed and constructed specifically for this experiment consists of a pair of vertical spindles separated by a distance of 21 in. and mounted in a grey cast iron housing 36″L×12″W×8″H with walls 1″ thick. See figure 2. Each spindle is mounted in a pair of ABEC class 9 angular contact bearings preloaded so as to remove radial play. The bearing seats were bored out on a Societé Genevoise jig boring machine after carefully aligning the casting on the machine bed with autocollimators; as a result the two axes are parallel to within 2 arc s. Rotation of each spindle is accomplished via a 24.25-in sine arm contained inside the cast iron housing. The ball mounted near the drive end of each arm bears against a $\frac{1}{2}$ in diam polished carbide flat mounted on a piezoelectric ceramic element; this PZT-carbide flat assembly is attached to the nonrotating spindle of a Boeckeler 2-in micrometer [9]. A spur gear mounted at the rear of the 3-in diameter micrometer barrel is engaged by a 2-in long pinion gear, which in

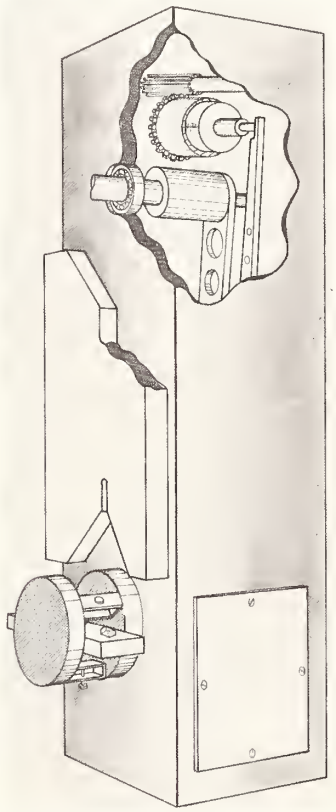


FIGURE 2. *Perspective drawing of the double crystal spectrometer.*

turn is driven through worm gearing by either of two motors, one for fast and one for slow motion.

Rotation of each spindle can be effected by three separate drives: (1) 0–1500 rpm motor-tachometers connected to the input worm shafts for fast motion; (2) 0–60 rpm Slo-Syn synchronous motors [10] driven by either a variable frequency quadrature oscillator or hand-rotated sine-cosine potentiometers also connected to the worm shafts for slow motion; (3) the PZT assemblies behind the ball-plane contacts excited by voltage bias for either very fine motions or as a means of applying a correction signal to the axis to close a servo loop for stabilizing the spectrometer against drift.

The full angular range of each axis is approximately $\pm 2.5°$; however, the rigid connection between the sine arm and its spindle can be unlocked so as to allow for free rotation of the spindle for alignment purposes. Each spindle is surmounted by a 6 in diam table, which in turn supports a 6 in diam cast iron crystal table by three legs of adjustable height. The diffracting crystals will be mounted upon these upper tables.

Rigidly attached to each crystal table is a cast iron arm of cross section 2 in$\times$2 in that supports a pair of glass corner cubes at a separation of 9 in symmetrically disposed with respect to the axis of rotation. Approximately 2 in in front of each corner cube is a high reflectivity flat attached to the spectrometer base. Each flat with its associated corner cube forms a 4 in Fabry-Perot cavity with a finesse of 20, the length of which changes as the spindle to which the corner cube is attached is rotated. This is the means by which the angular displacements of the spindles are measured. A pair of cavities are employed for each spindle so that eccentric motion of the axis does not contribute errors to the angular measurement.

Mounted upon a cast iron plate seated upon the top surface of the spectrometer casting in the region between the two spindles are the interferometer flats, various optical elements for directing 6328 Å laser beams into and away from the cavities, and photomultipliers to yield electronic fringe signals from the interferometers. Orthogonal eigen-polarizations are employed in each cavity as a means of sensing the direction of rotation.

The angular noise levels inherent in the spectrometer have been measured by rotating one of the spindles to such a position that one of its interferometers is at the half intensity point on a fringe. In this position, the associated photomultiplier output has a maximum sensitivity to rotation. We have found the inherent angular noise level of each axis to be less than 10^{-4} arc s; furthermore, one can accomplish controlled rotations of this amount. Thus the angular noise in the spectrometer is sufficiently low to ensure that the annihilation measurement can be carried out to 1 ppm.

5. Conclusion

In summary, we are planning a one ppm measurement of the Compton wavelength in a positron-electron annihilation experiment. It has been shown that such a measurement is equivalent to a determination of the fine structure constant with an uncertainty of 0.5 ppm. This measurement has been made possible because of the advent of perfect crystals as diffracting elements, and because their grating spacings can be determined by comparing the crystals to the crystal of a scanning x-ray interferometer. Doppler broadening of the source will be reduced by employing Zeeman mixing of positronium in a 5 K gaseous helium moderator. The double crystal spectrometer that has been fabricated for this experiment has been shown to be capable of generating angles smaller than 10^{-3} arc s.

6. References

[1] DuMond, J. W. M., Lind, D. A., and Watson, B. B., Phys. Rev. **75**, 1226 (1949).
[2] Knowles, J. W., Can. J. Phys. **40**, 257 (1962).
[3] Knowles, J. W., Proceedings of the Second International Conference on Nuclidic Masses, Vienna, 1963 (Springer-Verlag, 1964), pp. 113–129.
[4] Deslattes, R. D., these Proceedings; Curtis, I., et al., these Proceedings; Bonse, U., et al., these Proceedings.
[5] Hart, M., Proc. Roy. Soc. **A309**, 281 (1969).
[6] Deslattes, R. D., invited paper, these Proceedings.
[7] Sauder, W. C., J. Res. Nat. Bur. Stand. (U.S.), **71A** (Phys. and Chem.), No. 5, 347–353 (Sept.–Oct. 1967).
[8] Sauder, W. C., J. Res. Nat. Bur. Stand. (U.S.), **72A** (Phys. and Chem.), No. 1, 92–93 (Jan.–Feb. 1968).
[9] Arizona Tool and Die Company, 31 East Rillito Street, Tuscon, Arizona 85705.
[10] The Superior Electric Company, Bristol, Connecticut 06010.

DISCUSSION

B. N. Taylor: I would like to ask a general question of both you and Dick Deslattes. Can you give us an estimate of when the milennium is about to descend upon us?

P. H. M. Van Assche: No. (*Laughter*)

W. C. Sauder: That's a dangerous kind of answer to try to give. All I can say is what the present status of the spectrometer is. Its optics are finished. There is some testing of it going on, but in essence that part of the program is done. We're about to begin construction of the source. And that's about all I can say. Certainly by next summer at this time I'll go crazy if we don't have something out—like a photon. I want to see one photon. (*Laughter*) I don't believe all this, and I want to see it happen. (*Laughter*)

J. S. Thomsen: We now have a definite "maybe." (*Laughter*)

Optical Interferometry of the 220 Repeat Distance in a Silicon Crystal

R. D. Deslattes

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Repeat distances in certain highly perfect crystals are more sharply definable than are the wavelengths of x-ray lines. Such repeat distances can be measured in terms of optical wavelengths by coupling a separated plate version of the nondispersive x-ray interferometer of Bonse and Hart to an optical Fabry-Perot in a fashion permitting coordinated scanning. A rather primitive device intended originally for a feasibility study has yielded some preliminary results which are reported. Characteristic features of a second generation type device are also described.

Key words: Lasers; lattice parameter; x-ray/optical interferometry.

1. Introduction

Several years ago an impasse was reached in the optical testing of [1] a locally designed rather complex device fabricated for scanning x-ray interferometry. About that time I became aware of a rather simple gadget which J. L. Blayney [2], a graduate student of the late Hugo Benioff, had produced for calibration of a geodetic strain gage. Its operation appeared so self evident that there seemed no need for optical pretesting. It seemed that such a device could be of immediate value in x-ray interferometry.

A device similar to Blayney's though elaborated for adjustment purposes and equipped with a more refined drive than the original was quickly fabricated and an x-ray interferometer attached, and then sawed apart. Thereby began a search for x-ray interference signals which was to last over a year.

Following persistent failures in searches for x-ray interference, various diagnostic probes were added to the assembly. These were meant to look for vibration and/or drift or other reasons for failure to find x-ray fringes. Among other things employed was a unique field emission probe working at atmospheric pressures [3] which revealed an unpleasant combination of vibrationally active environment and vibrationally sensitive device. Finally, still as part of the diagnostic exercise, the elements of a primitive optical interferometer were added. These additions showed finally that, in a more quiet area by then available, drifts and noise were such that x-ray fringes should appear.

Fringes were then shortly obtained although a significant defect in the procedure for separating the parts of the x-ray interferometer had first to be remedied.

In retrospect, that original success might well have signaled an appropriate occasion to have redone the design to achieve a more integrated result. Alas, this was not done and the original artifact has remained in various degrees of operation and modification over the past 15 months [4].

The device is diagramed in figure 1. I shall first comment on its physical aspects including the optical interferometer and laser light source. Next I shall discuss the various operational procedures

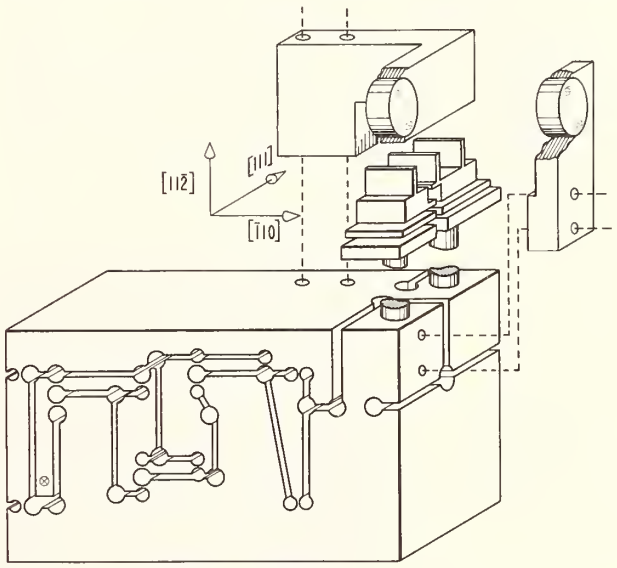


FIGURE 1. *Partial diagram of final stages of x-ray/optical interferometer.*
The driven point is labeled ⊗.

which have been worked out to establish alignment and permit relatively efficient measurement schedules. Finally, I shall report on performance limitations and numerical results obtained to date.

At the very outset it should be acknowledged that this device does not live up to the expectations one should have regarding the performance of this type

of interferometer. Experience with it has, however, been sufficiently good that one can see a finite group of modifications which should lead to a more acceptable second generation device [5].

2. Physical Description

Figure 1 shows an articulated brass monolith with surmounted x-ray and optical interferometers. It can be seen that there are two sizes of metal bending flexures provided at various places. Some have thin webs, about 0.1 mm to provide pivoting action, while others have 1 cm thick webs to provide great stiffness against which other, weaker springs (not shown) can effect microangular adjustments.

The outermost four thin flexures define a parallelogram which prescribes the trajectory of relative motion between the separate interferometer parts. The interior network of slots and thin flexures defines a three stage compound lever whose reduction factor is about 100. This system is driven through a toggle pin by a 10:1 lever actuated by a worm driven micrometer. The toggle pin contains piezoelectric disks which are used for ultramicro positioning in connection with servo operations described below.

The thicker flexures shown provide for weak spring driven adjustments for Bragg parallelism and Moire tilt. The weak springs are actuated by eccentric ball-bearing cams which are worm driven by motors located outside a closed metal shield.

The lasers used have contained either ^{20}Ne or ^{22}Ne plus ^{3}He. They are rf excited at 120 MHz. Stabilization is to the Lamb dip but the extent of the frequency modulation needs be carefully controlled since it can be comparable to the frequency sensitivity of the optical cavity.

The optical interferometer has a flat plate and a curved one both dielectric coated to reflectivities of 99.5–99.7 percent at 632.8 nm. They are placed quite close together (0.3 mm) to minimize requirements on laser stability and reduce pressure sensitivity. The whole is mounted in a sealed pressure vessel windowed for light and x-rays which is filled with helium under control of a density servo. The apparatus including laser source but excluding the x-ray source rests on a 2000 kg platform supported by pneumatic springs. This is in turn contained in a massive closed box. In operation, control is exercised from an adjoining room.

3. Alignment and Operation

There are three alignment exercises required to establish (a) the relative orientation of the separated parts of the x-ray interferometer, (b) parallelism of the flat mirror to the x-ray planes, (c) normalcy of the laser beam and adjustment of the optical cavity for high finesse.

a. X-ray alignment is secured in stages: the parts are separated after mounting to obtain a high degree of prealignment, then beams from an auxiliary x-ray source are bounced down the tunnel shown in figure 1, and finally, x-ray fringes are searched for and optimized. The tolerances on the three rotational degrees of freedom are nanoradians for tilt, microradians for Bragg parallelism and milliradians for the remaining rotation [6].

b. At first glance it might appear that the optieal propagation vector, direction of motion, and x-ray diffraction vector all require to be closely parallel. Fortunately, this is not the case, as it is easy to convince oneself that only the optical propagation direction and the x-ray diffraction vector need be closely parallel to one another [4]. Tolerance on this adjustment is about one milliradian per part per million.

The adjustment is carried out by the arrangement in figure 2. X rays from the source S are collimated (to about 2 arc min) by the pinholes P_1 and P_2.

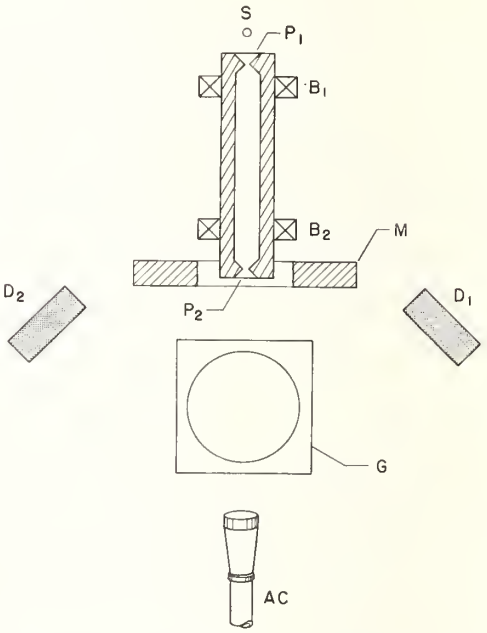


FIGURE 2. *Diagram of facility for making optical normal parallel to x-ray normal.*

These are mounted in bearings B_1 and B_2 so that their centration may be verified. This rotation also allows the adjustable mirror, M, in this way to have its normal brought to parallelism with the x-ray beam and the axis of the autocollimator, (AC). An accurate goniometer, G, is used to locate symmetric x-ray reflection (220, and $\bar{2}\bar{2}0$), with the aid of x-ray detectors, D_1 and D_2. When the table is brought to the mean of the two goniometer readings, it is known that the plane containing the diffraction vector is parallel to the optical axis. The flat mirror of the interferometer can be then adjusted until its normal also lies in a parallel plane. The operation repeated after rotation of the x-ray/optical interferometer by 90° about the x-ray beams suffices to secure the required parallelism between optical mirror and x-ray planes. It is readily carried out to

1 arc min and may be pushed to 10 arc s with the available apparatus.

c. The optical interferometer alignment is the easiest step but not quite trivial. As will be shown below, the working finesse of the curved mirror/flat mirror cavity is more than 600. This means that the tangent plane of the curved mirror at the beam center must be brought within a few seconds (the curved mirror radius is about 1 meter) of parallel with the flat mirror. This has been usually achieved by symmetrizing the off-axis modes in the optical cavity as viewed in a telescope.

4. Operational Characteristics and Performance Limits

In the past year's time, a number of learning experiences have transpired in this exercise. Each was associated with a particular mode of data taking and had the character of revealing one or more defects in the system and a reason for altering the mode of data taking. I shall detail a few of these since they suggest both the problems and the potential of this kind of measurement.

4.1. Initial Long Scan Data

In the early stages of this program I took it that the first task would be to find the integer component of the λ/d ratio. Measurements to this end were effected by long continuous scans. When the dual traces were analyzed they showed remarkable consistencies in one scan direction but significant differences compared to the reversed scan. The mean value of the integers were somewhat consistent though not always so. There were two problems here, one barometric, the other a drift between the two different interferometers.

To help eliminate the tedium (and mistakes) in manual counting of long recorder traces of optical and x-ray fringes, an electronic system was established to count x-ray fringes on the fly [7]. Negative-going zero crossings of each x-ray fringe were accumulated in a scaler and typed out at symmetric values of the 00l optical fringes. The large amount of data obtained in this way correlated most strongly with barometer effects and led to employment of a controlled atmosphere chamber.

On the other hand, no real confrontation with the drift problem came about in this way because of very long times required for each scan (ca 100 min for 1650 fringes). To get hold of this problem the system was arranged for rapid inter-optical fringe scan and slow scan through the region of the 00l optical maxima.

4.2. Short Scan Data

Arranging scan speeds of the order of 1 Å/min or so produced local scans of the sort shown in figure 3. Repetitions of such a scan at a particular optical fringe gave a measure of drift. This mode of operation, as also the previous one, exhibits a noise level reflecting the incoherent sum of correlated and uncorrelated disturbances of the two interferometers. Even with these limitations it was possible to obtain x-ray fringe fraction versus optical order data. These exhibited different slopes for different scan direction but the resultant fringe fraction mean appeared reasonable. The last statement was made easier to sustain by the large error which had to be assigned because of the presence of uncompensated, though

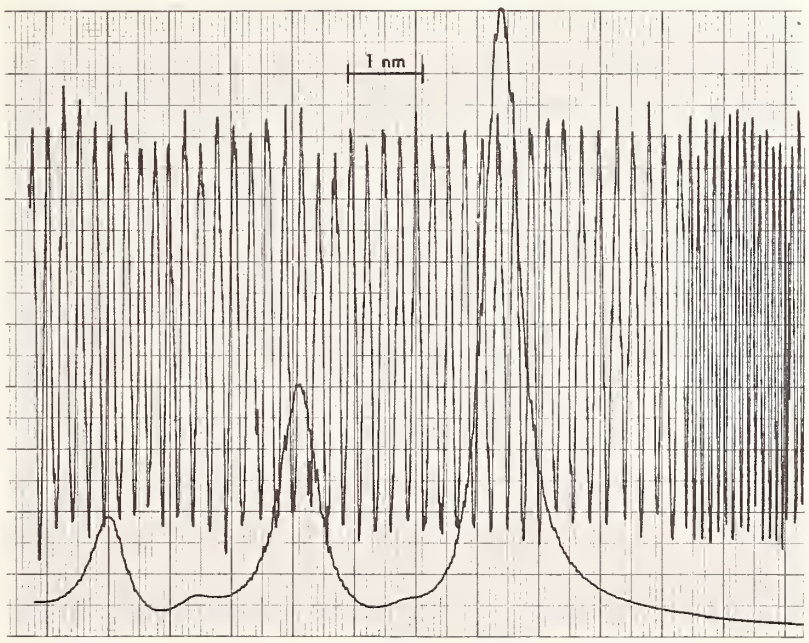


FIGURE 3. *Short scan through* 00l *optical maximum with nearby x-ray fringes shown.*

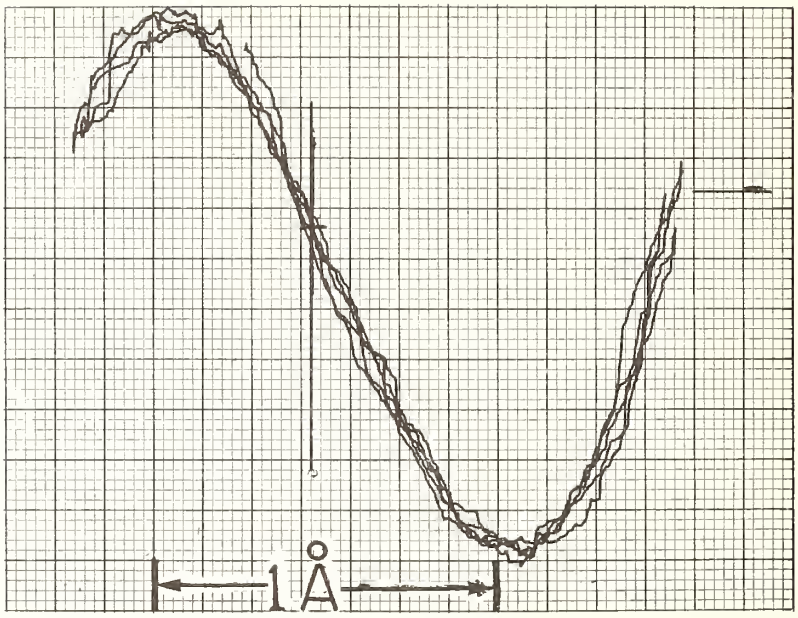


FIGURE 4. *Repetitive scans of x-ray fringe (ordinate) versus derivative of optical* 001 *maximum (abscissa).*
See text for technique.

correlated, noise. This same effect (namely, large noise) made it easier to accept that the two interferometers might be in uniform relative motion. Had this been the case then a simple average of the results obtained in opposing scan direction would have been appropriate.

4.3. Reduction of Correlated Noise

The first step in this direction was to impose a 0.1 Å sweep on the moving platform at about 20 Hz. Phase sensitive detection of the resulting optical modulation produced an approximate derivative signal. Against this one could plot the x-ray fringe signal on an x–y recorder as the device was slowly swept. This produced results such as shown in figure 4. Unfortunately, recorder dynamics enter the data and it takes a long time to produce this kind of phase measurement.

A significant further noise reduction follows using this optical derivative signal to close a feedback loop with the addition of a high voltage amplifier

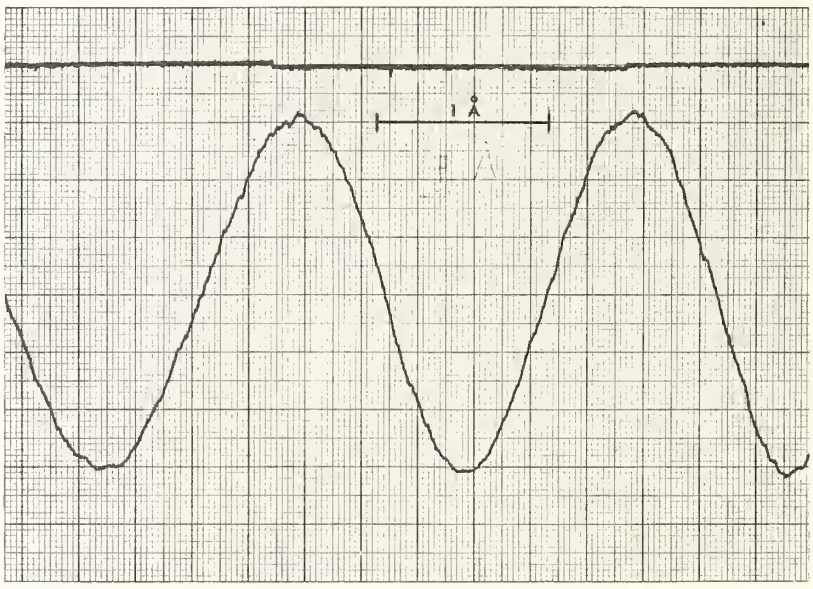


FIGURE 5. *Performance of system when locked to* 001 *optical maximum.*
X-ray signal shown corresponds to rapid drift shortly after closing box.

to drive the piezoelectric crystal. This gave results such as shown in figure 5. This shows thermal drift just after closing all available isolation barriers. It is evident that the positional noise level is about 0.01 Å. With this tool at hand, one begins to have the resolution needed to see rather subtle but serious problems in the system.

5. Results and Conclusions

It appears evident that operations of the required delicacy to produce a significant measurement are quite possible. The system which I have described suffers from serious inadequacy regarding the relative stability of the x-ray and optical parts. It is planned to use much harder material for the next generation device and to mount the x-ray and optical parts in a common silicon block.

In spite of these problems some information already emerges. In scanning only one optical fringe it is easy to obtain the x-ray fringe fraction to 5 percent or so. This already gives a value for the lattice parameter good to 30 ppm. It appears that this fraction is 0.3. The result is thus 1648.3 ± 0.1 lattice parameters per optical fringe at 25 °C.

The wavelength of the 633 nm line in $^{3}He^{20}Ne$ is taken to have the vacuum value of 632.9915 nm [2] and the index of refraction for 1 atm He at 25 °C taken as 1.000033 [9]. The implied value for d_{220} in the Si sample is 1.920070 ± 60 Å.

Previous measurements (with a completely independent optical adjustment) in air and depending on good weather for consistency give the result 1647.83 ± 0.05 lattice parameters per fringe. If this is multiplied by the ratio of the index of air (1.00029) to that of He, one obtains 1648.32 ± 0.05. This agreement cannot be regarded as other than satisfactory at the present stage of development.

This value is not yet significant which is probably a good thing. The true utility of this exercise will appear only when this is transferred first to x-ray wavelengths and then, independently of the wave lengths, to direct constants measurements.

6. References

[1] In these early phases my collaborators were Bowman, H., and Schoonover, R. M.
[2] Shopland, R. C., Bull. Seismol. Soc. Am. **56**, 337 (1960).
[3] Young, R. D., Rev. Sci. Instr. **37**, 275 (1966).
[4] Deslattes, R. D., Appl. Phys. Letters **15**, 386 (1969).
[5] This is being undertaken by J. A. Hammond.
[6] te Kaat, E., Z. Physik **243**, 14 (1971).
[7] I am indebted to C. Sidener for this electronics as well as for the remote control circuits.
[8] Mielenz, K. D., Nefflen, K. F., Gillilland, K. E., and Stephens, R. B., Appl. Phys. Letters **7**, 277 (1965).
[9] A.I.P. Handbook 6-95 (McGraw-Hill, New York, 1963).

DISCUSSION

J. S. Thomsen: If it's agreeable with the other authors, we will postpone the discussion of these three papers until the end of the session.

A New Determination of Avogadro's Number

I. Curtis and I. Morgan

National Physical Laboratory, Teddington, Middlesex, England

M. Hart[1] and A. D. Milne

Department of Physics, University of Bristol, Bristol, England

A new direct approach to the determination of the Avogadro Number using simultaneous optical and x-ray interferometry is described. The factors potentially limiting the accuracy are discussed in detail and preliminary results are presented.

Key words: Avogadro number; crystal perfection; silicon; x-ray interferometry.

1. Introduction

The value of the Avogadro number which is recommended by Taylor, Parker and Langenberg in their recent survey of fundamental constants [1] has been obtained indirectly through a least squares adjustment of some highly selected data. Accurate direct measurements, such as those by Bearden [2] and Henins and Bearden [3] do not determine an absolute value for the Avogadro number but rather the product of N and Λ^3, the Ångstrom-kilo x-unit conversion factor. This results from the use of x-ray crystallographic measurements of the lattice spacing of a crystal in x-units, a unit of length defined in terms of the wavelength of an x-ray emission line. As pointed out by Cohen and DuMond in 1965 [4], at least two inconsistent working definitions of the x-unit have been used, resulting in ambiguous interpretation of most of the early precision x-ray measurements.

Some of the inconsistencies have been eliminated with the introduction by Bearden of a newly defined unit—Å* in terms of which he has measured a complete range of x-ray wavelengths [5]. Unfortunately, this is also essentially an independently defined unit which is known no better with respect to the fundamental standard of length than the earlier x-unit. A conversion factor, in this case Λ^*, is still required to relate the x-ray scale to the absolute scale of length. The measurements of Bearden and Henins were recently used [1] to obtain accurate values of these conversion factors.

This unsatisfactory situation concerning our knowledge of the Avogadro number has been highlighted recently by the very large difference between the values recommended by the 1963 [4] and 1969 [1] least squares adjustments. The important difference between these two analyses is the inclusion in the second of data from a new source; the macroscopic quantum properties of superconducting materials. In view of the large change in value, which is many times the assigned standard error, it is highly desirable to obtain new independent data concerning any of the adjustable parameters of the least squares adjustment.

It is with this in mind that we have recently undertaken the measurement of the Avogadro number. Through the invention of x-ray interferometers [6, 7] a new method of measuring crystal lattice spacings, which does not depend on a knowledge of the x-ray wavelength, has become available. This will eliminate the necessity of conversion factors between x-ray and absolute measurements.

Briefly, our method is to measure absolutely the lattice parameter of a silicon sample by means of x-ray and optical interferometry. From our knowledge of the perfection of the silicon this can be directly related to the volume of the unit cell without loss of accuracy. Later, we will determine the density of our silicon crystals from direct measurements of the volume and mass of large pieces of single crystal. The lattice parameters of the density samples and the x-ray interferometer samples can be mapped and compared with great accuracy: double crystal topographs, obtained with high order Bragg reflections and short x-ray wavelengths can be used to examine large areas of crystal with a strain sensitivity $\delta d/d$ of 10^{-8} and a spatial resolution of a few micrometers [8]. Multiple Bragg reflection diffractometry, which has a sensitivity $\delta d/d$ of 10^{-9}, can be used to compare the lattice parameters of *different* samples and so can be used to transfer measurements from one sample to another [9]. Combining these results with a knowledge of the abundances and atomic weights of the three stable isotopes of silicon leads to a determination of Avogadro's number.

[1] At present, Senior Research Associate of the National Research Council at NASA, Electronics Research Center.

2. The X-Ray Interferometer Crystal

As the initial step in our measurement of the Avogadro number, we are making an absolute determination of the repeat distance of silicon in the [110] direction. Because our approach is quite different in detail from those adopted in other laboratories, we will describe our experimental arrangement at some length.

The method involves translating one crystal of an x-ray interferometer while simultaneously counting x-ray Moiré fringes and measuring the distance traveled with an optical interferometer. The scale of movements involved, however, demands exceedingly good mechanical stability of the various components. One x-ray Moiré fringe corresponds to a displacement of about 2×10^{-10} m, so that the relative position of the two parts of the x-ray interferometer must be controlled to approximately 10^{-12} m. Equally, to eliminate rotational components of the Moiré field, relative rotations must be less than 10^{-9} rads.

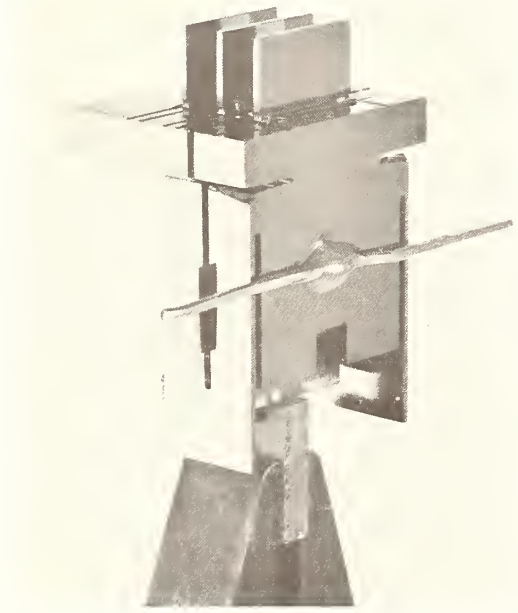


FIGURE 1. *Monolithic silicon scanning x-ray interferometer.* Approximately 3 cm high. The three x-ray wafers stand on two platforms which are strain isolated from the lower U-shaped springs by two horizontal saw cuts. The near edges of the platforms are polished to form mirrors for the optical interferometer. Rotational adjustments are made by applying couples to the short lever.

Although desirable, it is not possible to measure the relative movement of the x-ray wafers directly as they do not provide sufficient edge area for optical reflection. Optical reflectors which are remote from the position of x-ray measurement have to be employed. Any movement, either rotational or translational, of these reflectors relative to their respective x-ray interferometer wafers will introduce first order errors.

In order to obtain the required stability and in an attempt to eliminate sources of systematic error, we have constructed as many of the important components as possible from the same monolithic block of single crystal silicon. This approach also has the advantage that it is comparatively easy to control the relative orientations of the diffraction vector, and the optical reflectors. In the x-ray interferometer shown in figure 1, the three diffracting wafers are situated on two platforms, the larger of which carries two wafers and is rigidly mounted on a solid base. The third wafer is attached to the smaller platform and is connected to the rest of the interferometer by two thin strips of material which form the deformable springs of a simple spring strip traverse which has been folded about the midpoint of the springs. The construction of both the x-ray interferometer and the spring-strip traverse system from the same monolithic block of silicon endows the apparatus with high stability and insures the alinement of the diffracting wafers. Detailed design criteria have been discussed in an earlier report [10].

The surfaces, which have been cut with a high speed diamond saw, were oriented using the Bond method [11] and differ by less than 1 min arc from (110), ($1\bar{1}1$) and ($\bar{1}12$) planes respectively. Surface damage introduced by the machining was removed chemically with a nonselective etch.

Two adjacent vertical faces of the platforms are used as optical reflectors. Standard optical techniques have been used with a suitable chemico-abrasive polishing medium to give high quality surfaces without introducing a significant amount of damage into the bulk of the platform. In this way, the optical measurements can be made on the same piece of single crystal as the x-ray measurements. Because the orientation of the optical mirrors was controlled to within 1 min arc of the (110) planes of the crystal, optical interferometers which are sensitive only to the normal component of the mirror translation can, in our experiment, only measure the component of motion normal to the Bragg planes. The arrangement therefore *eliminates* systematic errors which might arise if the normal to the optical mirrors did not coincide with the traverse direction.

In all experimental arrangements which use elastic components in the design of the traversing system, there is a danger that the elastic strain will be communicated either to the x-ray wafers or to the region between them and the primary optical mirror. In our case we have isolated these important parts from the strained springs with two sawcuts just below the platforms (fig. 1). Nevertheless, the isolation may not be perfect. To measure the strains at any stage of displacement, Moiré topographs have to be taken of the whole area of the interferometer wafers. It is important to notice that in interferometers which have separate x-ray and optical mirrors, strains in the region between the two components may go undetected and cause first order systematic errors. As these weakly linked interferometers are sensitive to vibrations, a new technique has been

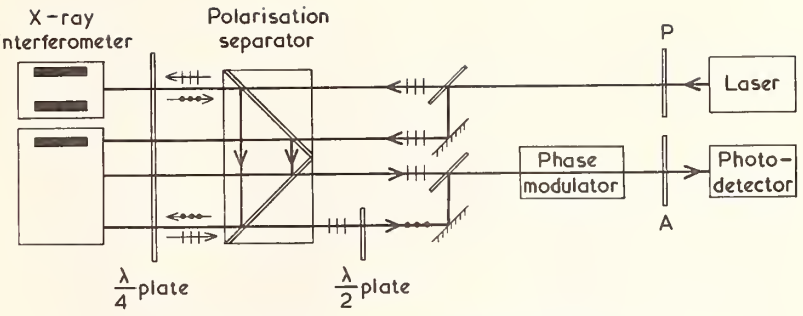


FIGURE 2. *Layout of the optical interferometer.*

developed for taking the Moiré topographs; the crystal is mounted on an antivibration platform and large areas can be studied by scanning both source and collimator across the crystal. A slit which is attached to the collimator collects the required diffracted beam which is recorded on a stationary film [12].

2.1. The Optical Interferometer

The optical interferometer which will be used to measure the relative displacements of the two platforms has been specially designed and contains some interesting features. In figure 2, the input beam from a He/Ne laser stabilized on the center of the Lamb-dip is linearly polarized before being split into two parallel components of similar intensity. These beams are transmitted through a polarization separator and an accurate quarter wave plate before each is reflected from a different platform at the x-ray interferometer. After a second traverse of the quarter wave plate, the vibration direction of the reflected beams is at right angles to the incident beams and they are internally reflected by the polarization separator. A second identical prism internally reflects the beams to produce two further beams which are parallel to the others. They are both reflected from the larger platform and after the double traverse of the quarter wave plate, are transmitted through the polarization separator. In order to prevent interference on recombining, the vibration direction of one of the beams is rotated through 90° by means of a half wave plate, and the beams are superposed by a full and a semireflector. Before passing through the analyzer, the beams are phase modulated at about 55 kHz so that the output can be accurately measured with a standard phase sensitive detector. The two superimposed beams are viewed with a photodetector through an analyzer whose transmission direction is at 90° to the resultant of the two vibration directions when they are in phase. That is, the output is a minimum when the phase difference between the beams is an integral number of half optical wavelengths. With this optical interferometer relative displacements of 0.2 A between the two reflectors can be measured. The four beams of the interferometer, which are spatially very compact, render it insensitive to parasitic bodily movements of the x-ray interferometer as the overall path length between the polarizing prisms and the reflectors remains unchanged. Second order effects are introduced due to the beams sampling different regions of the optical components. This sets high tolerances on the precision of the surfaces but these can be achieved with conventional techniques. The alignment of the optical interferometer with the [110] direction is a simple matter as we have ensured in the preparation of the reflecting surfaces that they deviate by less than 1′ arc from the (110) planes. With some of the components removed, the laser, which is positioned about three meters from the crystal, is adjusted until a simple reflection returns along the same path. Pulling effects of the optical interferometer on the laser can be reduced if necessary by a slight, but tolerable, amount of misalinement.

2.2. Apparatus

The mechanical drive to the spring strip crystal traverse is shown in figure 3. The x-ray and optical interferometers are mounted on a brass plate about 25 cm in diameter which is supported kinematically on a large block of lead. This is mounted on soft rubber mountings and forms a very good antivibration table whose natural period of vibration is about half a second. Surrounding the lead but isolated from it, is an accessory support ring which is set into the top of a concrete plinth. The ring and plinth carry the motor drives for the translation mechanism, the source of x-rays and the scintillation counter. There are also temperature enclosures (not shown) which surround the brass plate but do not include the x-ray tube housing or the electric motors. The immediate environment of the interferometers is therefore free of large sources of heat.

The only connections between the vibration free table and the concrete plinth are through weak couplings from the motors and flexible electrical leads. No detectable vibrations are transmitted through the motor couplings when the motors are stationary. Motor vibrations are quickly damped when the motor is stopped and do not upset the performance of the x-ray interferometer.

The translational motion of the x-ray wafer is produced by means of a lever and tangent arm

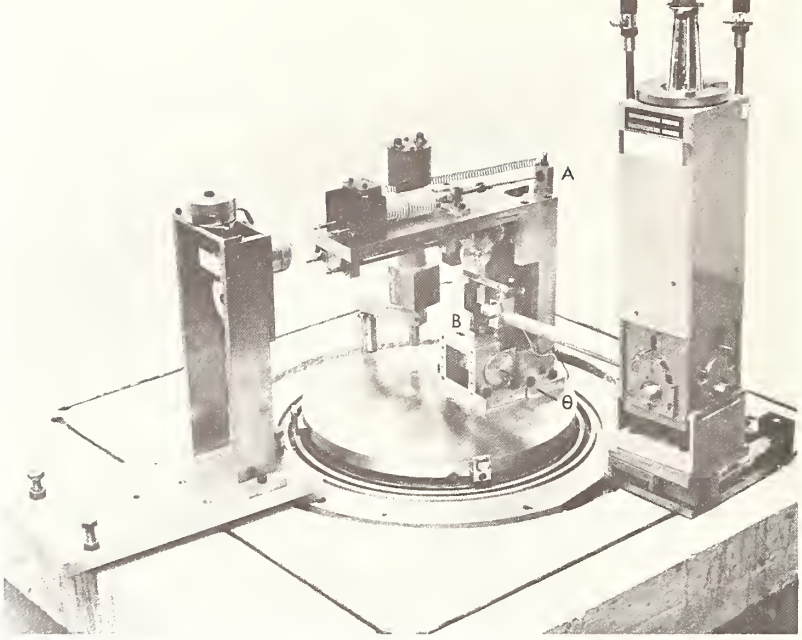


FIGURE 3. *General view of the x-ray apparatus.*

system. The robust lever A, 20 cm in length, is pivoted about an axis θ which is directly below and very close to the x-ray interferometer at B. The coupling to the interferometer is about 5 cm above the axis. A micrometer screw thread and wobble pin are used to move the lever and give a maximum range of about 8 mm displacement at the crystal. The micrometer is driven through a worm and wheel reduction of approximately 150:1 by either of two stepping motors which are separated by a 25:1 reduction gear. This provides a coarse adjustment which is necessary for scanning a large number of optical fringes in a reasonable time and a fine adjustment which is smooth enough to give individual x-ray Moiré fringes. The other two drives are used to control the relative orientation of the x-ray wafers during the experiment.

The details of the connection between the lever and the x-ray interferometer are shown in figure 4. A symmetrical leaf spring is mounted on the end of a small tube of piezoelectric ceramic material. The spring bears against a platelet of silicon which has been glued with epoxy resin onto the body of the monolith at the moveable side of the spring strip traverse. In a simple traverse the motion of the moveable platform is only closely rectilinear when the force is applied midway between the springs and equidistant from the platforms. With a folded traverse, the point of application of the force is not so obvious. In practice, however, it is found that when the force is applied between the two parts of the folded spring, a position can be found which gives a linear motion over a considerable range. As might be expected the height of the point of contact is extremely critical and allowance has been made to position the leaf spring with an accuracy of better than 1 μm by mounting the piezoelectric element on the cross-head of a fine-pitch parallel screw assembly. The purpose of the piezoelectric element is to give a much finer method of displacement than is convenient with the mechanical drive. It also has the advantage of being completely free of vibration and heat. The displacement it can produce is, however, limited by the depoling voltage of the ceramic and the present arrangement has a range of a few hundred x-ray fringes.

Since the leaf spring is mounted on a lever which rotates, the direction of the applied force slowly changes as the wafer is displaced. This tends to rotate the x-ray wafer and introduce a rotational

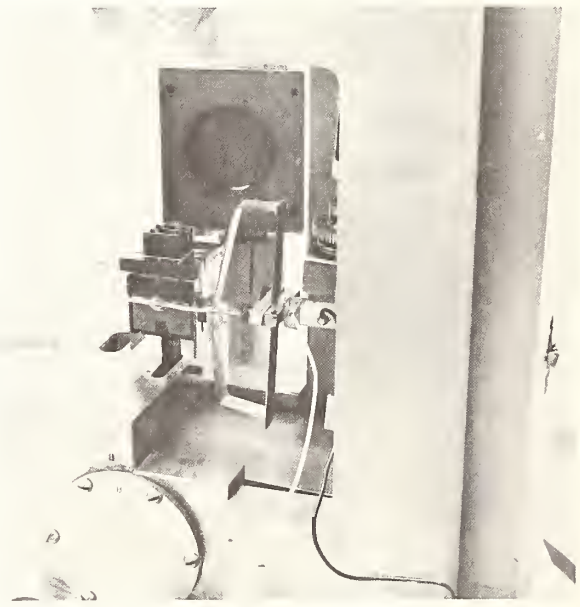


FIGURE 4. *Detail of the elastic scanning mechanism.*

component into the Moiré field. The counter, which is receiving an x-ray beam of a few mms in height from the interferometer begins to accept a larger fraction of a single fringe and, hence, the observed x-ray Moiré contrast is reduced. In order to prevent this effect, a couple of exactly the same strength as that introduced by the leaf spring is applied to the crystal. As a relative rotation of the two parts of the interferometer of 5×10^{-9} rads produces noticeable loss of contrast, quite delicate control is required. Here the monolithic construction is most helpful for one can achieve this remarkable control by simply extending a very weak spring or strip of elastic material which is attached to the torsion bar. Optimization of fringe contrast is normally required after about 15 optical fringes and can be achieved in a matter of seconds. Even when Moiré contrast is lost altogether during much longer scans, no difficulty is experienced in regaining the contrast quite quickly.

3. Experimental Progress

The apparatus has been used to obtain fringes over a range of more than 100 μm. Local scans of several hundred x-ray fringes at intervals of a few tens of optical fringes show no loss in the contrast that can be obtained at any position of the traverse. The elastically strained parts of the leaf spring tend to relax, which introduces a significant drift in the fringe pattern. Rather than being annoying this is a positive advantage as we do not have to supply any power during the actual comparison of x-ray and optical orders.

Our aim is to measure the Avogadro number with an error no greater than 1 ppm so that we have arbitrarily set 0.2 ppm as the tolerable error in each stage of the measurement. To achieve this in the lattice parameter measurement requires a relative displacement of 100 μm. In view of the rather large number of atomic planes in this distance it is not feasible to count each one individually. Fortunately, this is not necessary as the precision of our optical interferometry allows us to make a step-wise progression identifying the excess fraction at each stage. The accuracy of our measurement therefore increases as we make larger displacements.

In an experiment of this nature the environment must be carefully controlled as variations in temperature, pressure and humidity can introduce errors directly. The sensitive equipment which is mounted on the antivibration mounting is surrounded by two enclosures. The inner acts as a shield and rests directly on the brass base plate. The outer which fits into the accessory ring has thermostated water flowing through it and controls the absolute temperature of the apparatus. Two further polystyrene enclosures surround the concrete plinth to provide isolation from the slight variations of the ambient room temperature. These are always less than half a degree in any twelve-hour period. The absolute temperature of the silicon block will be measured by three copper constantan thermocouples in series which are embedded in silicon chips and suspended in close proximity to the x-ray interferometer. They will be calibrated against a platinum resistance thermometer. An error of 1/30 °C only introduces an error of 0.2 ppm in the lattice spacing.

4. Conclusions

Our philosophy throughout the design of the experiment has been to *eliminate* known sources of systematic error rather than to make corrections for them. This has led us, for example, to monolithic construction of the x-ray interferometer and the optical mirrors. Monolithic construction of the spring traverse system is also very convenient and while there have been some technical difficulties associated with this approach, the advantages obtained are overwhelming. If the experiment is to be successful, there is one fundamental condition which must be satisfied: the distance measured with the optical interferometer must be the same as the displacement normal to the Bragg planes of the x-ray wafer. With monolithic construction this condition can be easily fulfilled and, more important, can be verified throughout the experiment. There are two separate errors involved here. The optical interferometer must measure only the component of the translation normal to the Bragg planes—this we have insured—and there must be no change in separation between the x-ray wafers and their adjacent optical mirrors—we can obtain moire topographs of the whole region of interest at any position of the scanned range.

We have already obtained a sufficient range of displacement and x-ray Moiré fringes over the full range with the x-ray interferometer described herein. Our optical interferometer has achieved the required precision of measurement. Soon we will have determined the absolute lattice spacing of silicon with an accuracy of 0.2 ppm.

In our measurement of the density of silicon we are also adopting the same approach. After high precision materials characterization, a direct determination of the masses and volumes of a suitable batch of crystals will be made. The limiting factor in the final value of Avogadro's number is our present knowledge of the relative abundances of the three stable isotopes of silicon. Further work on this aspect of the determination is required.

5. References

[1] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[2] Bearden, J. A., Phys. Rev. **137**, B455 (1965).
[3] Henins, I., and Bearden, J. A., Phys. Rev. **135**, A890 (1964).
[4] Cohen, E. R., and Dumond, J. W. M., Rev. Mod. Phys. **37**, 537 (1965).
[5] Bearden, J. A., Rev. Mod. Phys. **39**, 78 (1967).
[6] Bonse, U., and Hart, M., Appl. Phys. Let. **6**, 155 (1965).
[7] Bonse, U., and Hart, M., Z. Phys. **188**, 154 (1965).
[8] Hart, M., Sci. Progr., Oxford **56**, 429 (1968).
[9] Hart, M., Proc. Roy. Soc. **A309**, 281 (1969).
[10] Hart, M., J. Phys. 1D, Ser. 2, 1405 (1968).
[11] Bond, W. L., J. Sci. Instr. **38**, 63 (1961).
[12] Milne, A. D., to be published.

Precision Lattice Parameter Measurement by X-Ray Interferometry

U. Bonse,* E. te Kaat,** and P. Spieker

Physikalisches Institut der Universität, Münster, 44 Münster, Germany

An apparatus applying the two crystal LLL-x-ray interferometer to the absolute measurement of the lattice spacing of a silicon crystal is described.

Key words: Precision absolute lattice parameter measurement; two crystal x-ray interferometers; x-ray moire fringes.

1. Introduction

The simultaneous measurement of x-ray and optical interference fringes as proposed some years ago [1] is in principle capable of furnishing a direct and very precise absolute value of the lattice parameter of a perfect crystal which can then be used as an absolute length standard in the Angström range.

A basic requirement to perform this measurement is to move a component crystal of an x-ray interferometer against the other component crystals in a direction normal to the reflecting planes very smoothly over a range of some hundred microns. During this motion relative rotations of the crystals should not exceed about 10^{-2} s of arc.

2. Apparatus

We describe here a design where two completely separate crystals are employed in transmission [2] one of which is mounted on a leaf spring type Δs-traverse illustrated in figure 1, by which the smooth movement normal to the net planes is accomplished. The other crystal is located on a Δx-traverse which allows equalization of the distance between mirror crystal and analyser crystal and between mirror crystal and beam splitter crystal in order to optimize fringe contrast.

The basic unit of either traverse is a leaf spring design TU which is explained in figure 2. The outer frame O of TU is bolted rigidly to the main base plate of the instrument. Its inner part 2 is only free to move parallel to the arrows indicated in the figure because of the special leaf spring arrangement. The Δs-traverse consists of two units TU linked together by a rigid frame the top beam of which can be seen in figure 1 and also figure 3. The Δs-traverse features less than 10^{-2} s of arc rotation over about 300 micrometer (μm) range. With no major change the range can be extended to about 600 μm.

The Δx-traverse is supported on one side by a unit TU and on the other simply by a ball bearing. Δx-shifts are accomplished by a micrometer screw via a lever acting on the traverse as shown in figure 1. For silicon 220 net planes Δx-adjustment to about 5 μm with CuKα radiation (20 μm with MoKα radiation) is sufficient in order to obtain good contrast fringes [3]. Δx-alignment is necessary only once at the beginning of measurement. All TU units are preloaded by coil springs to double their effective range.

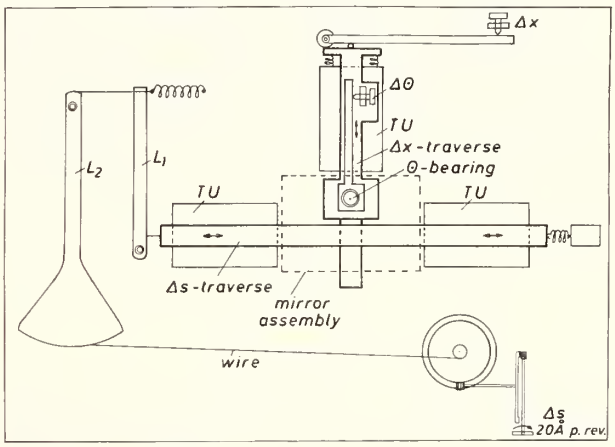


FIGURE 1. *Layout of combined x-ray and optical interferometer.*

The Δs-traverse is driven by a system of two levers with an overall reduction of about 1:400. The outer lever is rotated by winding a steel wire about 0.1 mm thick onto a shaft which is turned at another $1:4\times10^4$ reduction through a double worm and wheel assembly as illustrated in figure 1. One rotation of the input shaft corresponds to about 20 angströms travel of the Δs-traverse.

* Present address: Institute of Physics, University of Dortmund 46, Dortmund–Hombruch, Postfach 500.
** Present address: IBM Corporation, East Fishkill, N.Y.

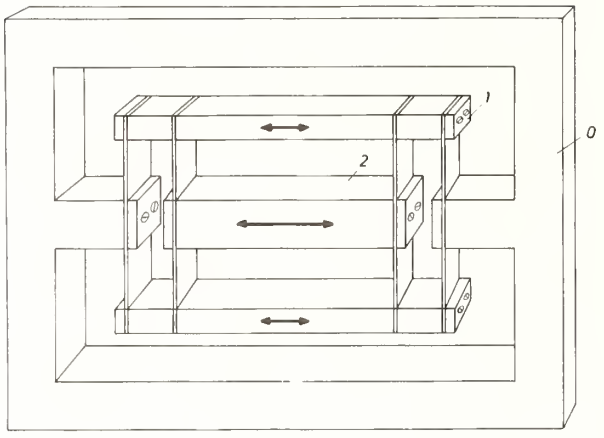


FIGURE 2. *Construction of traverse units TU used with Δs- and Δx-traverses.*

The wire is preloaded by an additional coil spring ensuring uniform wire tension and also eliminating mechanical play. Backlash as measured by reversing rotation of the input shaft is below 2 angströms. The input shaft is driven by a synchronous motor fed from either a variable frequency a-c supply or a hand-driven sin-cos potentiometer. In practice the traverse speed is limited by counting statistics of the x-ray fringe detecting system and not by mechanical limitations.

The light interferometer used is of the Fabry-Perot type. It is necessary to guarantee a rigid and stable coupling between crystal C_1 and FP-plate FP_1 and between crystal C_2 and FP-plate FP_2 and at the same time to have a provision for adjusting the reflecting plane of either crystal parallel to its conjugate FP-plate to about 30 s of arc. Figure 3 illustrates the construction of assemblies A of C_1 and FP_1 and B of C_2 and FP_2 respectively. By means of two differential screws with 50 microns effective pitch rotation about the ρ-axis (axis of rotational Moiré) and about the Θ-axis (axis about which the Bragg angle is aligned) can be performed between crystal and FP-plate in assembly A and B. The screws are preloaded by coil springs, and as a joint a steel pin is used, which is elastically bent for alignment. One revolution of the differential screw corresponds to about 200 s of arc tilt about either axis and the range is roughly 40 min of arc.

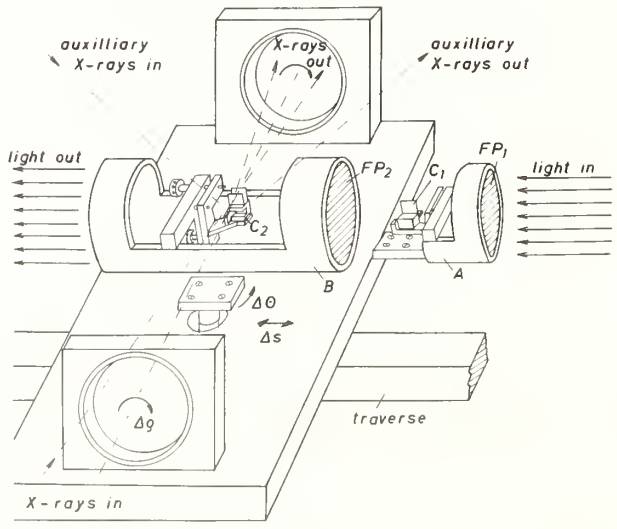


FIGURE 3. *Layout of assemblies of crystals C_1, C_2 with conjugate Fabry-Perot plates FP_1, FP_2 respectively and their provisions for ρ- and Θ-alignments.*

Assembly A is located inside B so that crystals C_1 and C_2 fit together with equal distances between the component crystals and also form a horizontal groove, one wall of which belongs to C_1 and the other to C_2. This groove is used for adjustment of the crystals with respect to each other about the ρ-axis by passing through it an auxiliary x-ray beam which is fivefold Bragg-reflected from the sides of the groove [2].

In the midpoint position of the Δs-range the distance between the FP-plates is about 1 mm. The whole instrument is placed inside a steel vessel which can be evacuated. At the plate distance stated a vacuum of 10^{-1} torr$\pm 10^{-1}$ torr is sufficient to use the vacuum wavelength value directly without troubling about corrections. It is not difficult to maintain such a poor vacuum even with the complicated interferometer inside the vacuum vessel.

As can be seen from figures 1 and 3 assembly A is bolted to the Δx-traverse with the provision of an additional Θ-rotation needed for relative Θ-adjustment between assemblies A and B as a whole. Correspondingly assembly B is connected with the Δs-traverse with an additional means for ρ-adjustment as illustrated in figure 3. ρ and Θ-adjustments are made by levers and micrometer screws with piezoelectric elements in between for remote fine adjustment.

The whole instrument including vacuum vessel rests on a simple inner tube type antivibration mount.

3. Adjustment Procedure

In a first step erystal C_1 is aligned parallel to plate FP_1 by means of a special double crystal technique illustrated in figure 4. Assembly A is placed on an alignment device consisting of a base plate B, a mounting plate M with ring R, and a mechanoelectric length transducer tr with lever 1 resting against the mirror side of FP_1. Mounting device and assembly A are positioned on the turntable of a commercial single crystal goniometer (Siemens) so that the goniometer axis is concentric with the figure axis of A as indicated in figure 4.

The mirror surface of FP is made normal to the goniometer axis by minimizing the variations in the transducer output while the turntable is rotated with the help of the screws S_ρ' and S_θ' which permit tilting M about the ρ and Θ-axes.

For aligning the net planes of C_1 an x-ray beam

reflected by C_1 and an auxiliary crystal C_H is used. C_H reflects by the same reflection order so that the rocking curve of C_1 and C_H is only about 3 to 4 s of arc wide. When C_1 is rotated about the goniometer axis the intensity intercepted by the counter is alined to least variations by turning screws S_ρ and S_θ of the assembly A which again facilitate adjustment about the ρ and Θ-axes. The accuracy of parallelism obtainable by this method gives less than 20 s of arc error which is sufficient to avoid an error in the d-spacing measurement due to nonparallelity of FP-plates and reflecting lattice planes.

Next assembly A is brought into its position inside B (see fig. 3) and FP_1 and FP_2 are aligned optically to parallelity. Finally by turning the differential

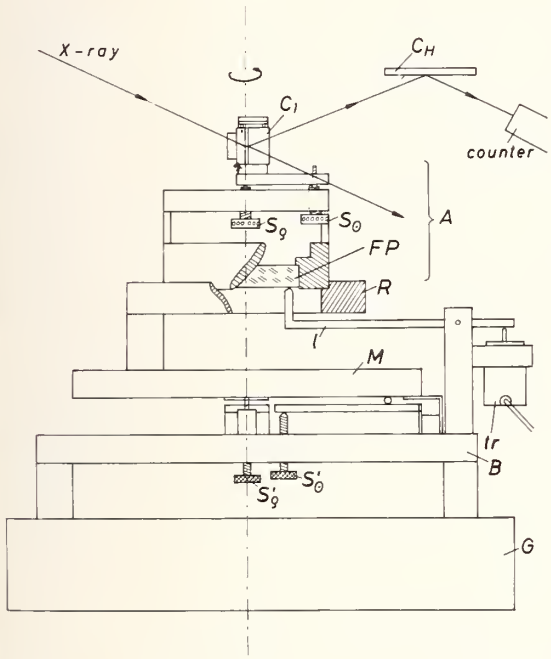


FIGURE 4. *Double crystal device for aligning parallel* C_1 *and* FP_1 *employing simultaneously x-ray reflectivity measurement and mechanical measurement of orientation of FP-plate.*

screws of assembly B crystal C_2 is made parallel to C_1 with the help of the main x-ray beam and the auxiliary x-ray beam as shown in figure 3.

4. Measurements

Since multiple beam interference and furthermore large FP-plates are used the optical measurement of a traverse shift of some hundred microns to an accuracy of 1 ppm or better and the control of truely rotationfree movement poses no particular problem, particularly if a stabilized and standardized laser is used. For the present measurements a Lamb-dip stabilized He–Ne gas laser is employed with an accuracy of at least 10^{-8}.

X-ray fringes are recorded by scintillation counters in both outgoing x-ray beams. Figure 5 gives an example of a sequence of typical x-ray fringes at two different traverse velocities of about 10 lattice spacings per minute (lower curve) and 20 spacings per minute (upper curve). Fringe contrast is slightly more and slightly less than 50 percent respectively. Fringes of this quality were obtained over the full range of the Δs-traverse corresponding to about 1.5×10^6 lattice spacings. With the use of MoKα radiation the intensity can be increased so that more spacings per minute can be recorded.

So far preliminary measurements have only been made. They show that the instrument is adequately stable and that alignment is possible to sufficient precision to measure the lattice spacing to at least an accuracy of 1 ppm. For the final measurement a cyclic traversing to increasing ranges is planned with iterative improvement of the d-value.

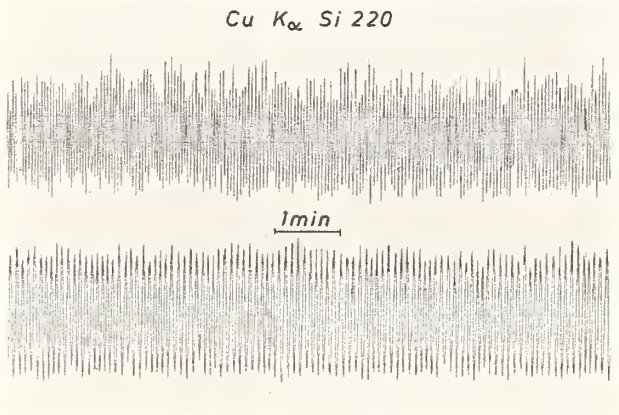


FIGURE 5. *Recorded x-ray fringes at speeds of about 20 (upper curve) and 10 (lower curve) net plane passages per minute obtainable at any position of the range of the* Δs*-traverse.*
Fringe contrast slightly above and below 50 percent respectively.

Since the preparation of the sample crystals is not very difficult other crystal materials of a variety of sources will be measured and their d-values and derived constants like the Avogadro number compared.

5. Acknowledgment

The Deutsche Forschungsgemeinschaft has provided financial support which is gratefully acknowledged.

6. References

[1] Bonse, U., and Hart, M., Z. Physik **188**, 154 (1965).
[2] Bonse, U., and te Kaat, E., Z. Physik **214**, 16 (1968).
[3] Bonse, U., and te Kaat, E., Z. Physik, to be published.

J. S. THOMSEN: We can now open discussion on the last three papers. Would the other two authors please come forward to receive questions. Perhaps I should first ask the respective authors if they have questions of each other that would be of general interest. (*Laughter*) No? Are there any questions from the audience then?

P. H. M. VAN ASSCHE: We realize now that the grating distance can be measured with absolute precision, but what do you think to be the utmost precision of the angular measurements if you are going to make wavelength measurements?

U. BONSE: What other measurements do you mean? I don't know who is asked.

R. D. DESLATTES: I guess me. As long as you're measuring Avogadro's constant, the game is over as soon as you measure the distance. If you want to get at anything else, then angle measurements become a problem. Angle measurements are hard. So is the x-ray diffraction theory. One can't believe that Bragg's law is right. Somebody has to work a little bit. But angle measurements to a hundredth of a second any place on the circle do not seem difficult. A part per million is easy. I don't know if that answers your question. A part in 10^8 is hard.

P. H. M. VAN ASSCHE: Yes, but then the problem can be to know how far you can trust Bragg's law.

R. D. DESLATTES: That's what I'd like to persuade Masao Kuriyama to teach me. Sticking with an optimum situation, symmetric Laue reflection in either a thin crystal or a thick crystal or optimally middle crystal, how far does one trust the Bragg relation? You have first-order immunity from index-of-refraction correction, but the whole notion of assigning a bulk index of refraction to a perfect crystal is just untenable.

J. S. THOMSEN: There is also the problem you mentioned earlier in your talk, Dick, that, even if the angle is measured to infinite accuracy and the Bragg law is confirmed to infinite accuracy, we still have this line width of the order of 300 parts per million.

R. D. DESLATTES: I don't think any of us are thinking about measuring x-ray lines except as a hobby. But I think Dr. Van Assche's question pertains to nuclear gamma rays where the normal full width at half maximum can be 2 to 5 parts per million or less.

J. S. THOMSEN: Then you're in a totally different ball park.

R. D. DESLATTES: Right. And that's the only place to really belabor the part per million business.

J. S. THOMSEN: But to do x-ray lines better than a part per million becomes really horrendous.

R. D. DESLATTES: Well, let me make a comment on that. X-ray lines were reference standards for measuring constants such as h/e, and so on. It was that fact that made it important to measure x-ray lines to a part per million. I don't believe that in their remaining application, which is to determine atomic energy levels, that a part per million is terribly relevant. I think we'll have to abandon that problem a little or at least I don't feel its urgency. Let's put it that way.

U. BONSE: I would like to ask the other author: I couldn't quite see on the slides how you mounted the optical plates to the crystals. I think that appeared to me a problem as far as I could see there.

R. D. DESLATTES: Right. It was a very bad problem in the original thing. But we don't intend to repeat those mistakes again. The answer is with as much glue as possible and as few mechanical joints as possible in the original. But in any future design the optical components and the x-ray components will be in the stomach of a big block of silicon even though they won't be a monolith *a la* Mike's. But it won't be a separated structure either. One tries to get everything that is possible unadjustable inside of a block of silicon. Epoxy is pretty good stuff.

M. HART: Could I make a comment on Dick's reticence about Bragg's law? The details of the reflectivity, the coherent reflectivity, of perfect crystals—that is, the shape of the reflection curve if you like to call it that—have been verified to about 0.1 percent. 0.1 percent of the width is in fact something like 1 part in 10^8 of the Bragg angle. So I really don't think we have any problems, at least in the x-ray wavelength range.

R. D. DESLATTES: I agree, but no one has put this into practical use. I agree as far as shape is concerned. The correct thing to do is to plug in the parameters to a dynamical theory and fit the profile observed and extract from that fit the wavelength parameter. That is vastly different from applying Bragg's law corrected by index of refraction to the peak or any other feature of an x-ray line. That's my only point. I would hope that (preferably) a young chap would undertake to reduce this to a simple-minded recipe so that old guys like me could use it. It's really a little bit much to do what I just said, namely, to turn on the full machinery of the dynamical theory every time you want to find out where you are. It ought to be possible to make a perturbation type correction from Bragg's law to a dynamical result.

M. HART: I believe the dynamical theory has been around since almost before x-ray diffraction, but not quite. Ewald didn't quite make it. I think it has been well verified.

R. D. DESLATTES: There's no question about that. There's no question of verification. It's a question of convenience and use and whether one actually uses it.

U. BONSE: With the computer there is no problem —once you understand it, I mean.

R. D. DESLATTES: I'll try.

C. BECKETT: What degree of perfection do you require in crystals in order to achieve this part per million accuracy in Avogadro's constant?

R. D. DESLATTES: Lots.

C. BECKETT: What degree of perfection?

R. D. Deslattes: Yes, lots.

C. Beckett: How do you determine that?

R. D. Deslattes: That's a problem. There's physical perfection. There's Stephan Peiser's crystal chemical perfection. There's isotope abundance. All of these things will require a great deal of work, and I'm sure it hasn't been done yet. But in principle the tools are available. It will be extremely difficult. The tools are the very enlightened study of defect clusters and things of that nature, high-resolution topography, isotope analysis. All of these things have to be applied, and it's going to cost a great deal of money to do that and a great deal of someone's time. But I don't think there's anything in principle wrong with the situation.

C. Beckett: Even if you achieve an ideal world of pure substances, unless you can grow these crystals at absolute zero they still will have defects.

R. D. Deslattes: Correct. But if you soak it at the right temperature where the equilibrium concentration of vacancies is tolerable but their mobility is still finite. If you soak it long enough, you can get to the point where the barrier is not at the part per million level for that story. It arises someplace further down.

C. Beckett: The fallacy there though is that the kinetics of the process have very much the same activation energies as the vacancy energies.

R. D. Deslattes: Right.

C. Beckett: And, in fact, you grow crystals at temperatures not too far from melting point.

R. D. Deslattes: Right. Well, I think that maybe what you are suggesting is that you will have to use in the Avogadro type of experiment a crystal that grows at room temperature, perhaps in solution, where these arguments do not limit it, but where we must face the problem of stoichiometry.

C. Beckett: Well, the point I'm really leading up to is I think you're going to have to determine the imperfections.

R. D. Deslattes: Agreed.

C. Beckett: And in the Physical Review Letters within the last week or two I saw a paper on the positron annihilation determination of vacancies in crystals. The details I didn't read carefully. But I think you're going to have to go into a great deal of depth in order to avoid this problem of imperfections in crystals.

R. D. Deslattes: You are undoubtedly correct.

U. Bonse: I would like to make a comment about the money which is involved. I have sometimes the impression that the fear of the difficulties involved here is much overdone. For instance, if you have to control an x-ray Moiré fringe pattern to keep the fringe distance constant, say at a separation of 2 millimeters or so, you're already controlling rotations of thousandths of seconds. How is this done? Not by big machinery. It's simply done by levers this long (indicating) and piezoelectric elements in between. So many people seem to believe that to do this you have to have much more machinery than you actually do. And they don't realize that this has been done in all these devices. Anyone who sees x-ray Moiré fringes is able to control thousandths and fractions of thousandths of seconds like any one of us does. And there are no millions of dollars involved, but fractions of that.

FINE AND HYPERFINE STRUCTURE IN SIMPLE ATOMS

No metal is perfect, it's sure,
So a Josephson junction's impure.
But Brodsky can fathom
The hydrogen atom,
So for alpha it's surely the cure.

HAROLD METCALF

Quantum Electrodynamic Theory: Its Relation to Precision Low Energy Experiments *

Stanley J. Brodsky[1]

Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14850
and
Stanford Linear Accelerator Center, Stanford University, Stanford, California[2] 94305

At present not even one of the crucial tests of quantum electrodynamics indicates any serious discrepancy. The key to the unraveling of previous conflicts has been the ac Josephson junction e/h determination of the fine structure constant α and new algebraic and numeric techniques for the evaluation of higher order radiative corrections—especially fourth order contributions to the Lamb shift and sixth order contributions to the electron and muon anomalous magnetic moments. In addition, recent measurements of the muon moment and the hyperfine splitting of muonium now permit a test of QED free of hadronic and weak interaction complications at a precision of 5 ppm.

Key words: Anomalous magnetic moments; fine structure constant; hyperfine splitting of hydrogen; hyperfine splitting of muonium; Lamb shift; quantum electrodynamics.

* Supported in part by the U.S. Atomic Energy Commission and the National Science Foundation.
[1] Avco Visiting Associate Professor.
[2] Permanent address.

1. Introduction

As far as we know, quantum electrodynamics provides a mathematically exact description of the electromagnetic properties of the electron and muon. To the extent that the electromagnetic properties of the nucleus are known and hadronic contributions to vacuum polarization are understood, the theory also provides the fundamental dynamical theory of the relativistic atom, including its external electromagnetic interactions.

At present quantum electrodynamics has reached the extraordinary state where not even one of its crucial tests indicates any serious discrepancy with its predictions. The key to the unraveling of previous conflicts of theory and experiment has been the independent determination of the fine structure constant α (via the ac Josephson junction measurements for e/h) and recent advances in algebraic and numeric computational techniques for the calculation of higher order radiative corrections—especially fourth order contributions to the Lamb shift and sixth order contributions to the anomalous magnetic moments of the electron and muon. In addition, recent measurements of the muon moment and the hyperfine splitting of muonium have now for the first time permitted a test of quantum electrodynamics free from complications of hadronic and weak interactions at a precision of 5 ppm.

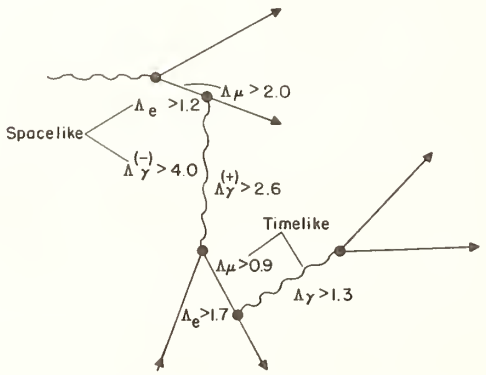


FIGURE 1. *Composite picture of high energy QED measurements.* The lower bounds for Λ (in GeV) correspond to 95 percent confidence limits on possible modifications of the photon and lepton space-like and time-like propagators. A review and references to individual experiments may be found in reference [1].

In this talk I will review only the most recent and relevant developments in the precision tests of quantum electrodynamics—those which have had a role in the resolution of the most perplexing discrepancies between theory and experiment and those which seem to point the way to an understanding of the fundamental basis and possible limits of the theory.

Before beginning this review, however, we should note that the high energy—high momentum transfer experiments are now in essentially perfect agreement with the theoretical predictions. The most beautiful tests of this type are those from colliding beams: e^+e^-, e^-e^- elastic scattering and $e^+e^-\to 2\gamma$, $e^+e^-\to\mu^+\mu^-$ annihilation, all of which are free from hadronic complications. The Born approximation structure of the theory—i.e., the validity of the Dirac current for the lepton and Maxwell's equations—has now been verified to small distances approaching 10^{-15} cm. The experimentally obtained lower bound [1] on possible high momentum transfer cutoffs for the various lepton and photon propagators are shown in figure 1.

Although high momentum transfer tests are essential for detecting possible new interactions or deviations at short distances, they are in practice only sensitive to Born diagrams. Tests of the high order corrections, including those involving renormalization effects require the very high precision atomic hyperfine and fine structure measurements and precise determinations of the electron and muon anomalous magnetic moments. As we shall see, the Lamb shift and hyperfine measurements are sensitive to the dynamical effects of quantum electrodynamics through fourth order in perturbation theory, as well as relativistic recoil corrections which emerge from the covariant treatment of the hydrogen atom bound state. The measurements of the magnetic moment of the electron are on the threshold of checking quantum electrodynamics through sixth order in perturbation theory. Besides this, at the level of precision now possible in studying the muon's $g-2$ value, one is able to probe the effect in an isolated electrodynamic system of very interesting hadronic and weak interaction contributions buried in the vacuum polarization; this leads to limits on the e^+e^- annihilation cross section into the entire spectrum of hadrons. In addition, statements about the polarizability of the proton structure itself can be inferred from the fantastically precise measurements of the ground state hyperfine splitting of hydrogen.

The validity of QED of course pertains to the validity of our understanding of all atomic physics, the analysis of the fundamental constants, and possibly the understanding of elementary particles. But perhaps the underlying goal of the precision tests is aesthetic: the hydrogen atom is the fundamental two-body system and perhaps the most important tool of physics; fifty-seven years after the Bohr theory the challenge is still there to calculate its properties to the highest accuracy possible.

2. Precision Tests of Quantum Electrodynamics

All of the precision tests of QED hinge on the value of the fine structure constant $\alpha=e^2/4\pi\hbar c$. Because of the precision measurements of e/h via the ac Josephson effect in superconductors and the massive reanalysis of the data relevant to the determination of the fundamental constants by Parker, Taylor, and Langenberg [2], α can now be determined to better than 2 ppm precision from experiments totally independent of QED input. The

least square adjusted result—which has now almost been canonized—is [3]

$$\alpha^{-1} = 137.03608(26) \qquad (1.9 \text{ ppm}). \qquad (1)$$

Thus finally the input constants necessary for comparing theory and experiment are known sufficiently to permit critical and unambiguous tests of theory.

2.1. The Anomalous Magnetic Moments

The classic and basic test of quantum electrodynamics[3] is the anomalous magnetic moment of the electron $a_e = (g-2)/2$. Thus far it has been one of the most stunning triumphs of theoretical analysis.

Very recently Wesley and Rich [4] have reported a determination of a_e to a precision of 6 ppm:

$$a_e^{\text{exp}} = 0.001\ 159\ 644(7). \qquad (2)$$

Although the new experiment is based on the same spin precession method used by Wilkinson and Crane [5], the reasons for the large difference with the final result [6] $a_e^{\text{exp}} = 0.001\ 159\ 549(30)$ for the older measurement is not understood. Promising new methods utilizing RF resonance techniques [7] or the change in flight time of ground state electrons in a magnetic field [8] have also been developed, although their precision is not yet comparable with the Wesley-Rich result.

Theoretically, the electron's interaction with the external electromagnetic field is completely specified by quantum electrodynamics. In general, contributions of order α^n to the anomalous moment are obtained in perturbation theory from Feynman diagrams for the free electron vertex containing n virtual photons, each diagram requiring up to $3n-2$ nontrivial parametric integrations. The diagrams which have been computed thus far are shown in figure 2.

The present theoretical prediction for the electron anomalous moment is

$$a_e^{\text{th}} = (\alpha/2\pi) - 0.32848(\alpha^2/\pi^2) + [0.26(5) + 0.13(\text{est})](\alpha^3/\pi^3). \qquad (3)$$

The first two terms are the famous second and fourth order results obtained by Schwinger [9] and by Sommerfield [10] and Petermann [11]. The sixth order coefficient consists of (a) the contribution $(0.055\cdots)$ for fourth order vacuum polarization contributions to the sixth order moment calculated analytically by Mignaco and Remiddi [12], (b) a contribution -0.154 ± 0.009 for second order vacuum polarization insertions in the fourth order vertex recently evaluated numerically by Kinoshita and myself [13], and (c) the contribution 0.36 ± 0.04 from photon-photon scattering diagrams evaluated numerically by Aldins, Brodsky, Dufner, and Kinoshita [14].

[3] The experiments can be idealized as a measurement of the static electron in isolation from other dynamics. [The contribution from hadronic vacuum polarization is smaller than even the muon contribution:

$$(\alpha^2/45\pi^2)(m_e^2/m_\mu^2) \sim 10^{-10}.]$$

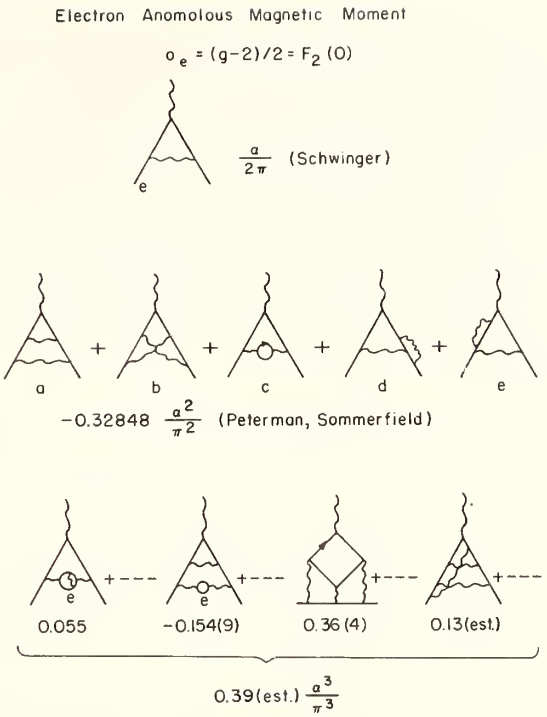


FIGURE 2. *Types of Feynman diagrams which contribute to the anomalous magnetic moment of the electron.*
Only representative diagrams are shown for the sixth order contributions. Mass, charge, and wave function renormalization counter terms are understood.

The graphs which have not been explicitly evaluated are the sixth order vertices with no electron loop insertions [there are 28 distinct types of diagrams of this type]. The sixth order coefficient includes a dispersion theory estimate [15, 16] $0.13(\alpha^3/\pi^3)$ for these three photon radiative corrections.

Although it does not eliminate the necessity for a full calculation of the sixth order moment the estimate of Drell and Pagels [15] and Parsons [16] strongly suggests that the final result for the non-electron loop graphs will be positive and numerically small.

The new result obtained by Wesley and Rich (with $\alpha^{-1} = 137.03608(26)$ is

$$a_e^{\text{exp}}(\text{W.R.}) = (\alpha/2\pi) - 0.32848(\alpha^2/\pi^2) + (0.54 \pm 0.58)(\alpha^3/\pi^2) \qquad (4)$$

which is very consistent with the present indicated sign and magnitude of the sixth order coefficient.

Further experiments and further development of the theoretical result[4] will be required before we can

[4] The exact calculation of sixth order radiative corrections to the lepton vertex is obviously a horrendous task. There are two central problems: (1) the reduction of matrix elements with three loop integrations to Feynman parametric form and (2) the multidimensional integration of the resulting integrand. In the photon-pboton scattering contribution calculation of reference [14] all the trace algebra and substitutions required to accomplish step (1) were performed automatically using an algebraic computation program written by Hearn [17]. The resulting 7-dimensional integration was performed numerically using a program originally developed by G. Sheppey which on successive iterations improves the Riemann integration grid through a random variable sampling technique.

be confident that QED is confirmed through sixth order in perturbation theory. The complete evaluation of the remaining graphs, though difficult, seems technically feasible with presently available algebraic and numerical computation techniques.

The muon anomalous moment is also an extremely valuable test of precision quantum electrodynamics. As a result of recent calculations, the difference of muon and electron anomalous moments has now been completely calculated through sixth order in quantum electrodynamics. The difference arises from the perturbation theory diagrams for the muon vertex with internal electron loops as shown in figure 3.

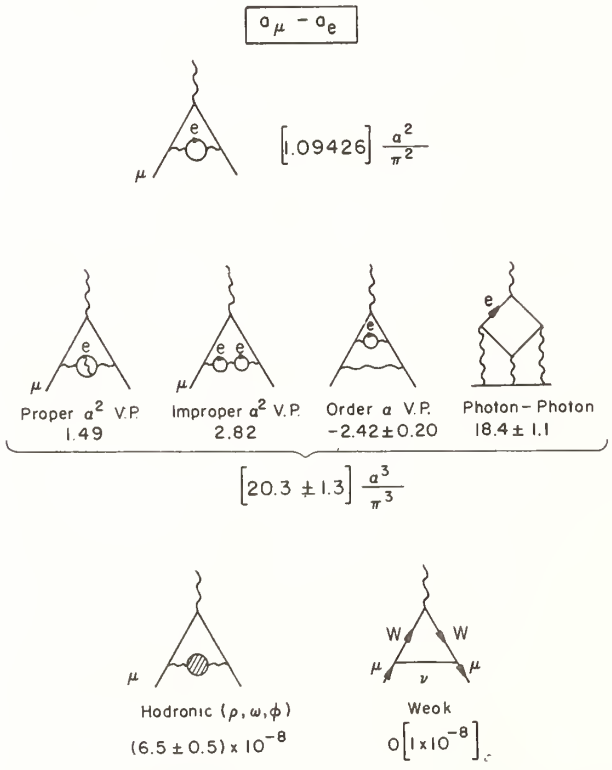


FIGURE 3. *The distinct types of Feynman diagrams which contribute to the difference of muon and electron anomalous moments.*
The contributions of muon and hadronic vacuum insertions to the electron vertex are negligible.

The complete QED result [18] is

$$a_\mu{}^{\rm th} - a_e{}^{\rm th} = 1.09426(\alpha^2/\pi^2) + [20.3 \pm 1.3](\alpha^3/\pi^3)$$
$$= 616(1)\times10^{-8} \qquad \text{(QED)} \qquad (5)$$

including the sixth order contributions from new analytic-numerical evaluations of second order vacuum polarization insertions [13, 19]; fourth order vacuum polarization [20, 21]; and photon-photon scattering contributions [14] due to the electron current. This result differs by $[-0.9\pm0.3](\alpha^3/\pi^3)$ from previous compilations [1, 2] which did not include the complete nonlogarithmic remainder of the second order vacuum polarization contributions. In addition, hadronic vacuum polarization calculated from the Orsay data for electron-positron annihilation in the ρ, ω, and ϕ regions gives the contribution [22]

$$\Delta a_\mu{}^{\rm th}(\text{hadronic}) = 6.5(5)\times10^{-8}. \qquad (6)$$

The result of the CERN measurement [23] for the anomalous moment of the muon is

$$a_\mu{}^{\rm exp} = 0.001\ 166\ 16(31). \qquad (7)$$

If we combine this with the Wesley-Rich result for $a_e{}^{\rm exp}$, we can check the theoretical result for the difference of muon and electron moments directly:

$$a_\mu{}^{\rm exp} - a_e{}^{\rm exp} = 652(32)\times10^{-8} \qquad (8)$$

$$a_\mu{}^{\rm th} - a_e{}^{\rm th} = 623(2)\times10^{-8}. \qquad (9)$$

The theoretical uncertainty does not take into account further (positive) contributions from hadronic vacuum polarization beyond the ϕ resonance region or possible weak interaction contributions which may be of order 1×10^{-8}. Nevertheless the one standard deviation agreement between theory and experiment is a remarkable success for the application of QED to the muon. Even at the present precision this agreement gives an interesting bound on the electron-positron annihilation cross section integrated over the entire hadronic spectrum [14]. Photon propagator cutoffs or negative metric photon [24] mass less than 5 GeV are ruled out to 90 percent confidence.

2.2. The Lamb Shift and Fine Structure of Hydrogenic Atoms

The historic tests of quantum electrodynamics have been the energy levels of the hydrogen atom [see fig. 4]. More recently the testing ground has been extended to other hydrogenic atoms and especially positronium and muonium for which the complications of hadron dynamics are remote. The dynamics of these atoms are completely specified

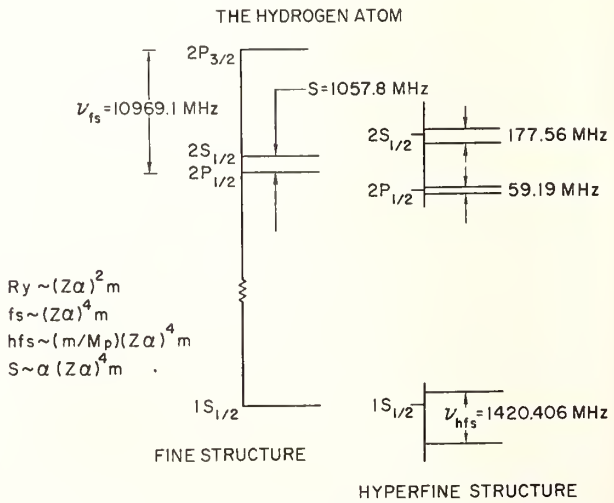


FIGURE 4. *The* n=1 *and* n=2 *levels of the hydrogen atom.*

by the interaction density of QED, $H_I = e\bar{\psi}\gamma_\mu\psi A^\mu$ plus the Maxwell and Dirac equations as expressed in perturbation theory in the form of the Feynman rules. The theoretical setting for the exact covariant treatment for the bound states of the hydrogenic atom is the Bethe-Salpeter equation [25]. If the atomic level experiments are idealized as photoabsorption measurements then as shown by Low [26] the line centers of the absorption spectra are determined by the eigenvalues of the full Bethe-Salpeter equation to at least order α^3.

The typical potential terms (irreducible kernels) which must be considered for the hydrogen atom

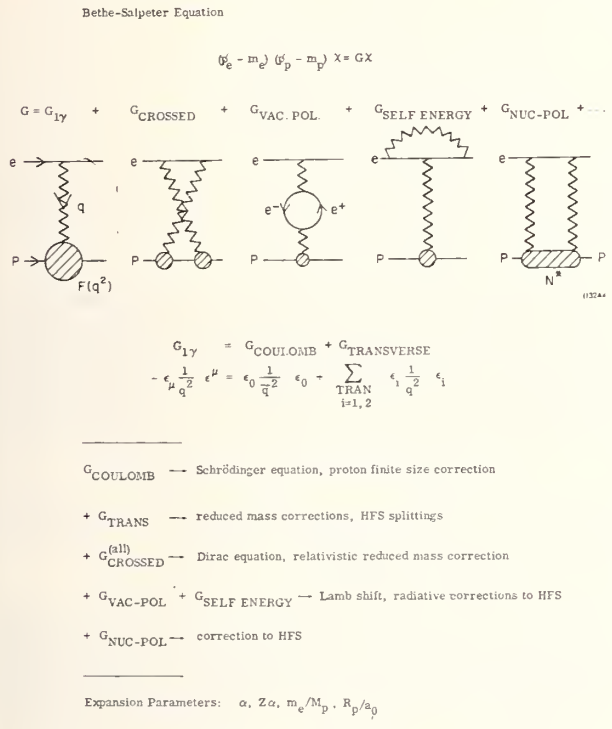


FIGURE 5. *Exact calculation of the hydrogen spectrum.* The typical kernels required for calculation of the energy levels of the H-atom to the present precision are shown. The one photon exchange contribution can be separated into Coulomb and transverse parts in the CM frame. The effects of strong interactions are summarized by form factors in $G_{1\gamma}$ plus nuclear polarization contributions as indicated in $G_{NUC\text{-}POL}$ The main effects of adding the higher order kernels are listed below the diagrams. The available small expansion parameters are also given. (From reference [27]).

are shown in figure 5. Of course the spectrum of the Bethe-Salpeter equation has not been solved exactly [28] and one must make recourse to the available small parameters. In particular, in the limit $m_e/m_p \to 0$, $R_p/a_0 \to 0$, $\alpha \to 0$ (i.e., the neglect of recoil, the proton finite charge distribution, and radiative corrections) one recovers the Dirac equation for an electron in a Coulomb field, with the famous Dirac-Sommerfeld degeneracy of the $S_{1/2}$ and $P_{1/2}$ levels. This degeneracy, however, is rather delicate and is removed by any modification of the Coulomb interaction. In particular, the s-state binding is strengthened by vacuum polarization and weakened by the proton finite size, by nonreduced mass recoil corrections and, most important, by the QED modification of the electron's charge and magnetic interactions with the proton. Over the years there has been considerable technical progress extending Bethe's historic calculation [29] of the contribution to Lamb interval ($2S_{1/2}$-$2P_{1/2}$) [from the one photon self-energy correction to the electron] to include terms from order $\alpha(Z\alpha)^4 m \log(Z\alpha)$ up to order $\alpha(Z\alpha)^6 m$ [30]. Also in the past two years Grotch and Yennie [31] have developed a very convenient effective potential method to handle the m_e/m_p corrections and have verified the previous calculations [32] of order

$$(Z\alpha)^5 m_e^2/m_p \quad \text{and} \quad (Z\alpha)^5 \log(Z\alpha) m_e^2/m_p.$$

At the ± 0.1 MHz precision level of the experiments, one is also sensitive to the two-photon (fourth order) QED correction. Recently Appelquist and myself [33], using algebraic and numeric computation techniques similar to those used for the higher order corrections to the anomalous magnetic moments [14], have obtained a new result for the fourth order contribution to the slope of the Dirac form factor of the electron—i.e., the order α^2 contribution to the effective mean square radius of the electron Dirac current distribution. The new result differs from the previous calculation [34] due to a change in overall sign and net numerical differences in the contribution of the nonlogarithmic remainder of two of the five distinct Feynman diagram contributions. The largest numerical discrepancy was due to the results for the fourth order "corner" diagram [the fourth diagram in the second line of fig. 2] to the free electron vertex. Our numerical result for this contribution has very recently been confirmed by de Rafael, Lautrup, and Peterman [35] at CERN and also by Barbieri, Mignaco, and Remiddi [36] who have obtained a completely analytic result for the crucial corner diagram.

The new results increase the $nS_{1/2}$-$nP_{1/2}$ separation in hydrogenic atoms by 0.35(7) MHz$\times[Z^4(2/n)^3]$ (3σ conf.) over previous compilations [30, 2]. A tabulation of the theoretical contributions for the Lamb interval in hydrogen is given in table 1.

A comparison with the most precise Lamb shift measurements is presented in figure 6 and table 2. I have rather arbitrarily taken $\pm 1/2$ of the limit of error (L.E.) to indicate the uncertainty in the theory. The recent measurements of the large interval $nP_{3/2}$-$nS_{1/2}$ are also included in the comparison with theory, utilizing the theoretical formula for the $nP_{3/2}$-$nP_{1/2}$ fine structure separation and $\alpha^{-1} = 137.03608(26)$.

The comparison of theory with experiment now shows quite satisfactory agreement because of the recent modification of the theoretical result for dF_1/dq^2 in fourth order. The remaining inconsistencies appear more as contradictions between the various experiments than with theory.

2.3. The Hyperfine Structure of Hydrogenic Atoms

Very exciting recent experimental developments have now made the hyperfine splitting in muonium

VARIOUS CONTRIBUTIONS TO THE LAMB SHIFT IN H (n = 2)

DESCRIPTION	ORDER	MAGNITUDE (MHz)
2^{nd}ORDER – SELF-ENERGY	$\alpha(Z\alpha)^4 m\{\log Z\alpha,\ 1\}$	1079.32 ± 0.02
2^{nd}ORDER – VAC. POL.	$\alpha(Z\alpha)^4 m$	- 27.13
2^{nd}ORDER – REMAINDER	$\alpha(Z\alpha)^5 m$	7.14
	$\alpha(Z\alpha)^6 m\{\log^2 Z\alpha, \log Z\alpha, 1\}$	- 0.38 ± 0.04
4^{th}ORDER – SELF-ENERGY	$\alpha^2(Z\alpha)^4 m \begin{cases} F_1'(0) \\ F_2(0) \end{cases}$	0.45 ± 0.07 - 0.10
	$\alpha^2(Z\alpha)^5 m$	± 0.02
4^{th}ORDER – VAC. POL.	$\alpha^2(Z\alpha)^4 m$	- 0.24
REDUCED MASS CORRECTIONS	$\alpha(Z\alpha)^4 \frac{m}{M} m\{\log Z\alpha, 1\}$	- 1.64
RECOIL	$(Z\alpha)^5 \frac{m}{M} m\{\log Z\alpha, 1\}$	0.36 ± 0.01
PROTON SIZE	$(Z\alpha)^4 (mR_N)^2 m$	0.13

$\mathscr{L} = \Delta E(2S_{\frac{1}{2}} - 2P_{\frac{1}{2}}) = 1057.91 \pm 0.16$ (L. E.)

$\alpha^{-1} = 137.03608(26)$

$\Delta E(2P_{\frac{3}{2}} - 2S_{\frac{1}{2}}) = 9911.12 \pm 0.22$ (L. E.)

$\Delta E(2P_{\frac{3}{2}} - 2P_{\frac{1}{2}}) = 10969.03 \pm 0.12$ (L. E.)

TABLE 1. (*From Brodsky and Appelquist, reference [33].*)
The previous result for the F_1' contribution to the $H(n=2)$ Lamb interval was 0.102 MHz. In view of the new analytic work by Barbieri, Mignaco, and Remiddi [36], this contribution can be further revised to 0.44±0.04 MHz.

(μ^+e^-) the most precise test of quantum electrodynamics.

Telegdi's group at the University of Chicago [37] has now reported the first direct measurement of the muon moment using the double-resonance method. Two ground state Zeeman transitions of muonium $\nu_1(F=1,\ m=1\leftrightarrow F=1,\ m=0)$ and $\nu_2(F=1,\ m=-1\leftrightarrow F=0,\ m=0)$ are measured at a field B_0 where the frequencies are to first order field independent: $\partial\nu_1/\partial B=\partial\nu_2/\partial B=0$. The ratio of sum to difference of the frequencies determines the muon moment μ_μ in units of the electron's magnetic moment in the atom. The results correspond to a determination of the muon moment to proton moment ratio:

$$\mu_\mu/\mu_p = 3.183337(14). \qquad (10)$$

A comparison of the muonium hfs measurement with theory is thus finally free of questions of possible chemical shifts in previous determinations of this ratio for muons stopped in water. Using their own best result for $\nu_{\mu e}^{\rm hfs}$ [38], the Chicago group obtains the value

$$\alpha^{-1} = 137.03568(33) \qquad [2.5\ \text{ppm}] \qquad (11)$$

which is in good agreement with the non-QED determination of Taylor et al., [2]:

$$\alpha^{-1} = 137.03602(26) \qquad [1.9\ \text{ppm}].$$

The paradoxical fact, however, is that the chemical shift discussed by Ruderman [39] (whereby muons in the intermolecular space experience ~15 ppm less diamagnetic shielding than the H_2O-bound protons)

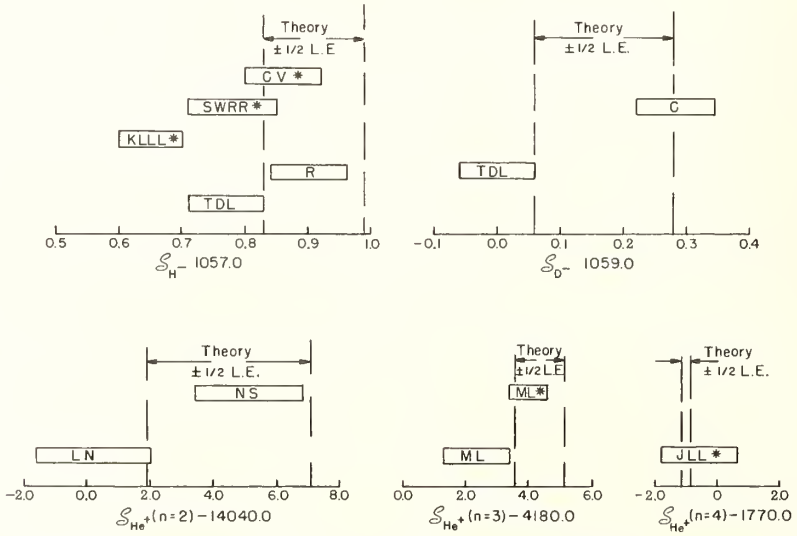


FIGURE 6. *Comparison of Lamb shift theory and experiment (see table 2).*

THE LAMB SHIFT IN HYDROCENIC ATOMS (in MHz)

	$\mathcal{S}_{exp}$ ($\pm 1\sigma$)	Theory ($\pm$ L.E.)	Exp-Th ($\pm 1\sigma$)
H (n = 2)		1057.91±0.16	
Triebwasser, Dayhoff, Lamb (a)	1057.77±0.06		-0.14±0.08
Robiscoe (b)	1057.90±0.06		-0.01±0.08
Kaufman, Lea, Leventhal, Lamb (c) [$(\Delta E\text{-}\mathcal{S})_{exp}$ = 9911.38 ±0.03]	(1057.65±0.05)		-0.26±0.07
Shyn, Williams, Robiscoe, Rebane (d) [$(\Delta E\text{-}\mathcal{S})_{exp}$ = 9911.25 ±0.06]	(1057.78±0.07)		-0.13±0.09
Cosens and Vorburger (e) [$(\Delta E\text{-}\mathcal{S})_{exp}$ = 9911.17 ±0.04]			-0.05±0.08
D (n = 2)		1059.17±0.22	
Triebwasser, Dayhoff, Lamb (f)	1059.00±0.06		-0.17±0.09
Cosens (g)	1059.28±0.06		+0.11±0.09
He^+ (n = 2)		14044.5±5.2	
Lipworth, Novick (h)	14040.2±1.8		-4.3 ±2.5
Narasimham, Strombotne (i)	14045.1±1.7		0.6 ±2.4
He^+ (n = 3)		4184.4±1.5	
Mader, Leventhal (j)	4182.4±1.0		-2.0 ±1.1
Mader, Leventhal (j) [$\Delta E\text{-}\mathcal{S}$ = 47843.8 ±0.5]	(4184.0±0.6)		-0.4 ±0.8
He^+ (n = 4)		1769.0±0.6	
Hatfield, Hughes (k)	1776.0±7.5		-3.0 ±7.5
Jacobs, Lea, Lamb (l)	1768.0±5.0		-1.0 ±5.0
Jacobs, Lea, Lamb (l) [$\Delta E\text{-}\mathcal{S}$ = 20179.7 ±1.2]	(1769.4±1.2)		0.4 ±1.3
Li^{++} (n = 2)		62771.0±50.0	
Fan, Carcia-Munoz, Sellin (m)	63031.0±327.0		260.0 ±333.0

TABLE 2. *Comparison of Lamb shift experiments and theory.* (From reference [33]).

[a] S. Triebwasser, E. S. Dayhoff, and W. E. Lamb, Jr., Phys. Rev. **89**, 98 (1953).
[b] R. T. Robiscoe and T. W. Shyn, Pbys. Rev. Letters **24**, 559 (1970); R. Robiscoe, Phys. Rev. **168**, 4 (0968).
[c] S. L. Kaufman, W. E. Lamb, Jr., K. R. Lea, and M. Leventhal, Phys. Rev. Letters **22**, 507 (0969).
[d] T. W. Sbyn, W. L. Williams, R. T. Robiscoe, and T. Rebane, Phys. Rev. Letters **22**, 1273 (1969).
[e] T. V. Vorburger and B. L. Cosens, Phys. Rev. Letters **23**, 1273 (1969).
[f] S. Triebwasser, E. S. Dayhoff, and W. E. Lamb, Jr., Phys. Rev. **89**, 98 (1953).
[g] B. L. Cosens, Phys. Rev. **173**, 49 (1968); see also T. V. Vorburger and B. L. Cosens, tbese Proceedings.
[h] E. Lipworth and R. Novick, Phys. Rev. **108**, 1434 (0957).
[i] M. A. Narasimbam and R. L. Strombotne, these Proceedings.
[j] D. Mader and M. Leventhal, International Conf. on Atomic Physics, New York University (1968).
[k] L. L. Hatfield and R. N. Hughes, Phys. Rev. **156**, 102 (1967); see also reference [2].
[l] R. R. Jacobs, K. R. Lea, and W. E. Lamb, Jr., Bull. Amer. Phys. Soc. **14**, 525 (1969).
[m] C. Y. Fan, M. Garcia-Munoz, and I. A. Sellin, Phys. Rev. **161**, 6 (1967).
Note: Tbe one standard deviation error limits assigned by reference [2] are used here for those experiments for which only "limits of error" are given.

and which was of great concern in interpreting the μ_μ'/μ_p' measurements in water is actually not applicable.

In a recent work, a group from the University of Washington and LRL [40] has measured the μ_μ'/μ_p' ratio in various chemical environments to the extraordinary precision of 2.5 ppm. As in earlier measurements [41] the proton resonance ω_p frequency and the muon (decay asymmetry) precession frequency ω_μ of stopped polarized positive muons are measured in the same magnetic field. The direct results are [42]

$$\mu_\mu'/\mu_p' = \omega_\mu/\omega_p = 3.183350(8) \quad H_2O$$
$$= 3.183355(8) \quad \text{NaOH solution.} \quad (12)$$

If the Ruderman model for magnetic shielding were applicable, then the muon frequency in the NaOH solution would be expected to be ~15 ppm lower than in H_2O—since the positive muon would rapidly combine with the NaOH and suffer about the same shielding as a proton. In fact it is 1.6 ppm higher. The correct picture, as discussed by the University of Washington/LRL group [40], seems to be that the slowed muons are first neutralized by capturing an electron. The "hot" muonium atom is then combined into neutral molecules with the substitution of the muon for a proton. The shielding of the muon in this situation is then very similar (within ~2 ppm) as the proton. In this new picture, the muon is rarely found in the intermolecular weak shielding region.

The final corrected value for the muon/proton moment ratio from the Univ. of Washington meas-

urements is

$$\mu_\mu/\mu_p = 3.183347(9) \qquad (2.8 \text{ ppm}) \qquad (13)$$

in excellent agreement with the Chicago group's result.

If we use this value and $\alpha^{-1} = 137.03608(26)$, then the predicted value for the muonium hyperfine splitting is [1, 31, 43]

$$\nu_{\mu e}{}^{\text{th}} = 4463.289(19) \text{ MHz} \qquad [4.3 \text{ ppm}]. \qquad (14)$$

This is very close to the weighted average of the two most recent results

$$\nu_{\mu e}{}^{\text{exp}} = 4463.317(21) \text{ MHz} \qquad \text{(Chicago)} \qquad (15)$$

$$\nu_{\mu e}{}^{\text{exp}} = 4463.249(31) \text{ MHz} \qquad \text{(Yale)}. \qquad (16)$$

Thus μ/e universality and quantum electrodynamics, free from hadronic effects, is now being tested at close to the 5 ppm level. The nuclear recoil corrections [44]

$$\delta_\mu = -(3\alpha/\pi)(m_e/m_\mu)\log(m_\mu/m_e)$$
$$= -179.7 \text{ ppm} \qquad (17)$$

which emerge from the Bethe-Salpeter covariant formalism are being checked to ~3 percent accuracy. The radiative corrections have been evaluated up to the 1 ppm level [43]. Finally we note that the new measurements [40] determine the muon to electron mass ratio to 3 ppm:

$$m_\mu/m_e = 206.7683(6).$$

The theory of the hyperfine splitting of the ground state of atomic hydrogen is also in a state of quite satisfactory agreement [1, 31, 43]:

$$(\nu_{\text{H}}{}^{\text{exp}} - \nu_{\text{H}}{}^{\text{th}})/\nu_{\text{H}}{}^{\text{exp}} = 2.5 \pm 4.0 \text{ ppm} - \delta_{\text{N}}{}^{\text{pol}}. \qquad (18)$$

The results are consistent with a small proton polarization correction [45].

The theory of the positronium ground state splitting has now been extended to terms of order $\alpha^6 m \log\alpha^{-1}$ by a remarkable calculation of Fulton, Owen, and Repko [46]:

$$\nu_{e^+e^-}{}^{\text{th}} = \alpha^2 \text{Ry}_\infty [\tfrac{7}{6} - (\alpha/\pi)(\tfrac{16}{9} + \log 2)$$
$$+ \tfrac{3}{4}\alpha^2 \log\alpha^{-1} + 0(\alpha^2)]$$
$$= 2.03415 \times 10^5 \text{ MHz} + 0(\alpha^6 m) \qquad (19)$$

using $\alpha^{-1} = 137.03608(26)$. The most recent experimental value [47] is

$$\nu_{e^+e^-}{}^{\text{exp}} = 2.03403(12) \times 10^5 \text{ MHz} \qquad [50 \text{ ppm}]. \qquad (20)$$

A calculation of the terms of order α^6 will be necessary for comparison with new experiments in progress at Yale which should attain a precision of 10 ppm [48].

3. Implications for Fundamental Physics

Quantum electrodynamics has never been more successful in its confrontation with experiment than it is now. This becomes especially apparent when one compares the values of α which can be derived to 5 ppm or better from the QED tests with its canonical non-QED value [see fig. 7 and table 3]. In addition to the values derived from the muonium and hydrogen hfs (assuming $\delta_p = 0 \pm 5$ ppm), α can be determined to high precision from the various level crossing measurements of the hydrogen fine structure separation [49]. Further work, both experimental and theoretical for the muonium hfs [50], positronium hfs, the fine structure of atomic helium [51], and the anomalous moment of the electron can conceivably produce values for the fine structure to a precision of 1 ppm. It is also possible that further high energy spin-analyzed electron-proton elastic and inelastic scattering data will eventually lead to an unambiguous determination of the nuclear polarization contribution to the hydrogen hyperfine splitting [52]. Measurements of the hyperfine splitting or Lamb interval in muonic-hydrogen $(\mu^- - p)$ would be of incredible interest for solving these hadronic problems [53].

A further essential element of our understanding of the hydrogen atom concerns its interaction with external fields. The general theory of the electro-

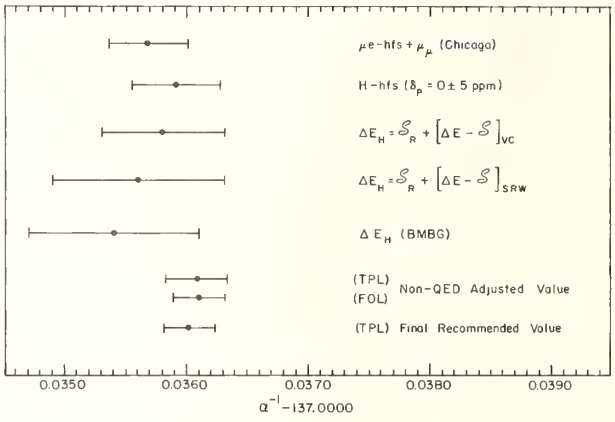


FIGURE 7. *Recent determinations of the fine structure constant.* The value for α obtained using the high precision "atomic bottle" measurement [54] of the $2P_{3/2}$–$2S_{1/2}$ interval depends critically on the value taken for the Lamb interval and is not included here. (See reference [2]).

Determination of α^{-1} to 5 ppm or better

TPL	Final Recommended Value (1969)	137.03602(21)	1.5 ppm
TPL	WQED Least Square Value (1969)	137.03608(26)	1.9 ppm
FDL	WQED Least Square Value (1970)	137.03610(22)	1.6 ppm
H-hfs	($\delta_p = 0 \pm 5$ ppm, 1969)	137.03591(35)	2.6 ppm
μ_μ + μ-e hfs	(Chicago, 1970)	137.03568(33)	2.5 ppm
ΔE_H	(BMBG, 1970)	137.0354(7)	5 ppm
H-fs	[$\mathcal{S}_R$ + $(\Delta E - \mathcal{S})_{SWR}$, 1970]	137.0356(7)	5 ppm
H-fs	[$\mathcal{S}_R$ + $(\Delta E - \mathcal{S})_{VC}$, 1969]	137.0358(5)	4 ppm

TABLE 3. *Recent determinations of the fine structure constant.* The value for α obtained using the high precision "atomic bottle" measurement [54] of the $2P_{3/2}$–$2S_{1/2}$ interval depends critically on the value taken for the Lamb interval and is not included here. (See reference [2]).

magnetic interaction of relativistic weakly bound composite systems is given within the Bethe-Salpeter formalism [55, 56]. Of greatest current interest is the general reliability of the Zeeman theory [56] (which is now required to 1 ppm accuracy) and the specific question of finite nuclear mass, radiative, and binding corrections to the electron's g_J value.

The basic formulae for these problems were given by Lamb in his famous papers [57]. More recently Grotch [58] and Hegstrom [59] have computed higher order corrections to the electron's g_J value. The dependence on nuclear mass is reflected in the ratio of g_J values for atomic hydrogen and deuterium. The prediction is

$$g(\mathrm{H})/g(\mathrm{D}) = \left(1+\tfrac{1}{2}(Z\alpha)^2\frac{m_e}{m_p}\right)\Big/\left(1+\tfrac{1}{2}(Z\alpha)^2\frac{m_e}{m_D}\right)$$
$$= 1+7.3\times10^{-9}\ [58] \qquad (21)$$

compared to the measured values:

$$\frac{g(\mathrm{H})}{g(\mathrm{D})} = 1+(9.4\pm1.4)\times10^{-9}$$
(Larsen, Valberg, Ramsey [60])
$$= 1+(7.3\pm3.0)\times10^{-9}$$
(Robinson and Hughes [61]) (22)

The radiative-binding corrections of order $\alpha(Z\alpha)^2$ to the g_J value can be tested by measuring the ratio in H to He^+ [58].

Quantum electrodynamics is the main link between the physics of the atom and elementary particle physics. In atomic physics it is the fundamental theory and the basis of all calculations. In particle physics and field theory it serves as the model and guide to the weak and strong interactions [62]. Despite its successes and despite the essential simplicity of its equations, it is clear that we are still uncomfortable with the theory as it is. For one thing, it is difficult to accept the infinite renormalization procedure as an essential part of a physical theory.[5] Ingenious extensions of the theory, especially the introduction by Lee and Wick of negative metric photons and leptons [24], and the possible nonpolynomial modifications due to gravitational effects, as proposed by Salam and Strathdee [64], can lead to finite physical theories.

As we discussed above, the comparison of the muon anomalous moment with conventional theory already rules out (to 90% conf.) negative metric photons with mass less than 5 GeV. The gravitational modifications of the theory seem to have no direct tests, but are interesting because the electromagnetic self-mass of the electron can, in principle, be calculated from the gravitational constant in such an approach.

From a second point of view, it is frustrating to have a theory which—as far as we know—provides an exact mathematical description of the physical world and yet tells us nothing about so many fundamental questions, especially the origin of charge quantization [66], the numerical value of α, and the problems understanding the existence of the muon and the symmetry of its interactions with those of the electron.

Despite these fundamental problems, the successes of quantum electrodynamics are phenomenal. Perhaps the most dramatic evidence of its validity is the electron's g-factor. Considering that there is no *a priori* reason for the bare lepton to have a Dirac moment ($g=2$), theory and experiment for $g=2(1+a_e)$ can be said to agree to nine significant figures.

The calculation of the entire sixth order contribution to the anomalous moment will almost certainly be completed within a few years—probably to within a numerical precision of $\pm0.1\alpha^3/\pi^3$. The necessity for measurement of a_e to a comparable precision will then be critical. The electron anomalous moment is perhaps the most precious and unique precision test of quantum electrodynamics; it is the only way we have to check the theory—and the correctness of the Taylor expansion in powers in α—through sixth order in perturbation theory.

In order to earry out this program, the value of α will be required to a precision near ±0.5 ppm. Thus the first order of business is to push the fundamental constants and precision tests as hard as possible, especially measurements of the muonium hfs, the hydrogen fs, and even the helium fs and positronium hfs.

An additional dividend of a muonium hfs measurement at this precision will be the determination of the proton polarization in the H hfs to $\sim$1 ppm; this could well be of fundamental importance to hadron physics [45]. Additionally, it should be emphasized that a high precision determination of the difference of muon and electron anomalous moments will yield invaluable information on the total contribution of the hadronic current to the vacuum polarization as well as a limit on the magnitude of the weak interaction correction to the electromagnetic current of the muon.

Although the comparison of theory with the Lamb shift measurements now shows satisfactory agreement, it is clear that further work is needed to improve the precision of the experiments and theory. Measurements in medium and high Z hydrogenie atoms could be of great value in checking out the various components and Z-dependence of the total theoretical result.

4. References

[1] Brodsky, S. J., and Drell, S. D., Ann. Rev. Nucl. Science (1970, to be published). Further details on the present status of quantum electrodynamics are given in this and the following reference.

[2] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[5] The infinities may, of course, only be symptomatic of an incorrect or asymptotic expansion. A convergent expansion in α may take a Padé form, for example [63]. On the other hand, in some mathematical field theory models studied by Jaffe and Glimm [64] the renormalization constants are infinite with or without a perturbation expansion. The best experimental clue we have to the possibilities of an asymptotic expansion is the test at sixth order of the electron anomalous moment.

[3] An even more precise value of e/h (to 0.46 ppm) has recently been obtained by Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970). This leads to $\alpha^{-1}=137.03610(22)$ [1.6 ppm].
[4] Wesley, J. C., and Rich, A., Phys. Rev. Letters **24**, 1320 (1970).
[5] Wilkinson, D. T., and Crane, H. R., Phys. Rev. **130**, 852 (1963).
[6] The recalculated result given here is obtained from analyses by Rich, A., Phys. Rev. Letters **20**, 967, 1221 (1968); Henry, G. R., and Silver, J. F., Phys. Rev. **180**, 1262 (1969); and Farley, F. J. M., Cargese Lectures in Physics, edited by M. Levy (Gordon and Breach, New York, 1968), Vol. 2.
[7] Gräff, G., Major, F. G., Roeder, R. W. H., and Werth, G., Phys. Rev. Letters **21**, 340 (1968); Dehmelt, A. (unpublished).
[8] Kincaid, B., Fairbank, W. M., and Knight, L. V. (unpublished).
[9] Schwinger, J., Phys. Rev. **73**, 416 (1948), **76**, 790 (1949).
[10] Sommerfield, C. M., Phys. Rev. **107**, 328 (1957); Ann. Phys. (N.Y.) **5**, 26 (1958).
[11] Petermann, A., Helv. Phys. Acta **30**, 407 (1957).
[12] Mignaco, J. A., and Remiddi, E., Nuovo Cimento **60A**, 519 (1969).
[13] Brodsky, S. J., and Kinoshita, T., submitted to the XVth Intl. Conf. on High Energy Physics, Kiev (1970).
[14] Aldins, J., Kinoshita, T., Brodsky, S. J., and Dufner, A., Phys. Rev. Letters **23**, 441 (1969), and Phys. Rev. (to be published).
[15] Drell, S. D., and Pagels, H. R., Phys. Rev. **140**, B397 (1965).
[16] Parsons, R. G., Phys. Rev. **168**, 1562 (1968).
[17] Hearn, A. C., Stanford Univ. Report ITP-247 (1969).
[18] The fourth order contribution to the muon-electron moment difference is the result obtained by: Surau, H., and Wichmann, E. H., Phys. Rev. **105**, 1930 (1957); Petermann, A., Phys. Rev. **105**, 1931 (1957); Fortschr. Physik **6**, 505 (1958); Elend, H. H., Phys. Letters **20**, 682, **21**, 720 (1966); Erickson, G. W., and Liu, H. (unpublished).
[19] The contribution of second order vacuum polarization insertions to the sixth order muon anomalous moment has also been calculated by de Rafael, E., Lautrup, B. E., and Peterman, A. (private communication). The results agree with those of reference [13].
[20] Lautrup, B. E., and de Rafael, E., Phys. Rev. **174**, 1835 (1968).
[21] Lautrup, B. E., and de Rafael, E., CERN preprint TH.1042 (1969).
[22] Gourdin, M., and de Rafael, E., Nuclear Phys. **10B**, 667 (1969).
[23] Bailey, J., Bartl, W., von Bochman, G., Brown, R. C. A., Farley, F. J. M., Jöstlein, H., Picasso, E., and Williams, R. W., Phys. Letters **28B**, 287 (1968).
[24] Lee, T. D., and Wick, G. C., Nuclear Phys. **39**, 209 [1969]; Lee, T. D., talk given at Topical Conf. on Weak Interactions, CERN (1969). Bailey, J., and Picasso, E. (to be published in Nuclear Phys.) have summarized the limits which can be obtained on various speculative theories from the agreement of theory and experiment for $a\mu$.
[25] Salpeter, E. E., and Bethe, H. A., Phys. Rev. **84**, 1232 (1951); Salpeter, E. E., Phys. Rev. **87**, 328 (1952); Gell-Mann, M., and Low, F., Phys. Rev. **84**, 350 (1951); Mandelstam, S., Proc. Roy. Soc. (London) **A238**, 248 (1952).
[26] Low, F., Phys. Rev. **88**, 53 (1952).
[27] Brodsky, S. J., and Primack, J. R., Ann. Phys. (N.Y.) **52**, 315 (1969); Phys. Rev. **174**, 2071 (1968).
[28] Relativistic two-body bound state spectra in quantum electrodynamics have recently been derived by Brezin, E., Itzikson, C., and Zinn-Justin, J., (to be published, Phys. Rev.). The proposed formulae, which are based on an analytic continuation of the eikonal forward scattering amplitude only take into account the long-range Coulomb potential and do not properly incorporate relativistic recoil corrections. The resulting spectra are incorrect beyond order $(Z\alpha)^4$ (S. J. Brodsky and D. R. Yennie, unpublished).
[29] Bethe, H. A., Phys. Rev. **72**, 339 (1947).
[30] Erickson, G. W., and Yennie, D. R., Ann. Phys. (N.Y.) **35**, 271, 447 (1965). These papers give a complete systematic evaluation of the order α radiative corrections to the Lamb shift.
[31] Grotch, H., and Yennie, D. R., Rev. Mod. Phys. **41**, 350 (1969).
[32] Salpeter, E. E., Phys. Rev. **87**, 328 (1952).
[33] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970), and Phys. Rev., to be published.
[34] Soto, M. F., Jr., Phys. Rev. Letters **17**, 1153 (1966). See also Weneser, J., Bersohn, R., and Kroll, N. M., Phys. Rev. **91**, 1257 (1953).
[35] de Rafael, E., Lautrup, B., and Peterman, A., CERN preprint Th.1140, 1970 (to be published).
[36] Barbieri, R., Mignaco, J. A., and Remiddi, E., Univ. of Pisa preprint, 1970. A complete analytic calculation of the fourth order vertex is now in progress (E. Remiddi, private communication).
[37] de Voe, R., McIntyre, P. M., Magnon, A., Stowell, D. Y., Swanson, R. A., and Telegdi, V. L. (submitted to the International Conference on Precision Measurement and Fundamental Constants, National Bureau of Standards).
[38] Ehrlich, R. D., Hofer, H., Magnon, A., Stowell, D., Swanson, R. A., and Telegdi, V. L., Univ. of Chicago Preprint EFINS-69-71 (1969).
[39] Ruderman, M. A., Phys. Rev. Letters **17**, 794 (1966).
[40] Hague, J. F., Rothberg, J. E., Schenk, A., Williams, D. L., Williams, R. W., and Crowe, K. M. (submitted to the International Conference on Precision Measurement and Fundamental Constants, National Bureau of Standards).
[41] Hutchinson, D. P., Menes, J., Shapiro, A., and Patlack, A. M., Phys. Rev. **131**, 1351 (1963).
[42] Another recent measurement of the ratio in water is

$$\mu_\mu'/\mu_p'=3.183362(30)$$

[Hutchinson, D. P., Larsen, F. L., Shoen, N. C., Sober, D. I., and Kanofsky, A. S., Phys. Rev. Letters **24**, 1254 (1970)].
[43] Newcomb, W., and Salpeter, E. E., Phys. Rev. **97**, 1146 (1955); Arnowitt, R., Phys. Rev. **92**, 1002 (1953).
[45] Theoretical work on the magnitude of the proton polarization correction is reviewed in references [1 and 31]. The contribution of nuclear polarization to the hfs of H [see fig. 5] is an extremely important subject since its calculation involves, in principle, most of the uncertainties of proton dynamics. For the hfs we are interested in spin-dependent dynamical quantities (whose spin-independent but isotopic-dependent parts are required for the calculations of the n-p mass difference, a program of limited success thus far). These current matrix elements also enter the study of inelastic e-p scattering which is currently of great interest and under intense study.
[46] Fulton, T., Owen, D. A., and Repko, W. W., Phys. Rev. Letters **24**, 1035 (1970), to be published.
[47] Theriot, E. D., Jr., Beers, R. H., Hughes, V. W., and Ziock, K. O. H., Phys. Rev. (to be published).
[48] Carlson, F. R., Hughes, V. W., Jr., and Theriot, E. D., Jr. (abstract submitted to the International Conference on Precision Measurement and Fundamental Constants, National Bureau of Standards).
[49] Metcalf, H., Brandenberger, J. R., and Baird, J. C., Phys. Rev. Letters **21**, 165 (1968), and references b, c, and d of table 2.
[50] A new experiment is being designed at Yale to determine both $\nu_{\mu e}{}^{\rm hfs}$ and μ_μ to a precision between 0.1 and 0.5 ppm [Crane, P., Hughes, V. W., zuPutlitz, G., and Thomson, P. A. (abstract submitted to the International Conference on Precision Measurement and Fundamental Constants, National Bureau of Standards)].

[51] The large $J=0$, $J=1$ interval in the 2^3P state of helium has been measured to 2 ppm. When calculations of the $0(\alpha^4$ Ry) contribution are completed, it should be possible to determine α to 1 ppm. See Hughes, V. W., Johnson, C. E., Kponoy, A., Lewis, S. A., and Pichanick, F. M. J. (abstract submitted to the International Conference on Precision Measurement and Fundamental Constants, National Bureau of Standards).
[52] Cherniak, V. L., Faustov, R. N., Zinojev, G. M., and Struminski, B. V., Dubna preprints (1969).
[53] Di Giacomo, A., Nuclear Phys. **B11**, 411 (1969).
[54] Reference e of table 2.
[55] Mandelstam, S., Proc. Roy. Soc. (London) **A238**, 248 (1955).
[56] Brodsky, S. J., and Primack, J. R., Ann. Phys. (N.Y.) **52**, 315 (1969); Phys. Rev. **174**, 2071 (1968); Brodsky, S. J., and Parsons, R. G., Phys. Rev. **163**, 134, **176**, 423 (1967).
[57] Lamb, W. E., Jr., Phys. Rev. **85**, 259 (1952); Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **79**, 549 (1950), **81**, 222 (1951), **86**, 1014 (1952).
[58] Grotch, H., Phys. Rev. Letters **24**, 39 (1970), and to be published.
[59] Hegstrom, R. A., Phys. Rev. **184**, 17 (1969), and to be published.
[60] Larsen, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23**, 1369 (1969).
[61] Hughes, W. M., and Robinson, H. G., Phys. Rev. Letters **23**, 1209 (1969).
[62] For a striking example see Cheng, H., and Wu, T. T., Phys. Rev. Letters **24**, 1456 (1970).
[63] Chisholm, R. (private communication).
[64] Jaffe, A., and Glimm, J., Commun. Math. Phys. **11**, 9 (1968).
[65] Salam, A., and Strathdee, J. (to be published).
[66] See especially Schwinger, J., Phys. Rev. **173**, 1536 (1968).

DISCUSSION

M. Leventhal: I was getting pretty depressed listening to your talk. (*Laughter*). Before you wiped out all the discrepancies, a group of us experimentalists had embarked on high Z Lamb shift measurements. With the excellent agreement between the theory and experiment now, is there any real point in extending these measurements up to high Z? Can you explain in more detail what we're actually testing, what aspects of the theory, in a language a simple-minded experimentalist can understand?

S. J. Brodsky: I think it's very important that one goes on and measures the Lamb shift for high Z hydrogenic atoms as well as possible. What we're testing is essentially the radiative corrections to the electron line—especially the one-photon self-energy correction—the fact that the electron's electromagnetic distribution is changed by one photon emission and absorption. The incredible thing is that we have in the Lamb interval the ability to change the radiative corrections since they're extraordinarly influenced by the binding. That's exhibited by the famous Bethe sum-over-states term, for example. Without binding the charge radius of the electron is even infrared divergent, but with binding it becomes infrared finite and thus essentially dependent on the binding corrections. So this is our opportunity to check the charge structure of the electron in quantum electrodynamics and see if it really works in a detailed manner.

Note that from the anomalous moment we only get one number, the static value of Pauli form factor. In the Lamb shift and also from the radiative corrections to the hyperfine splitting in hydrogen and muonium, we involve the entire distribution of the charge and magnetic moment due to radiative corrections.

Now, I'm not completely satisfied right now with the present status of the Lamb shift theory and experiment. The theory needs more work, and in addition the experiments certainly need improvement, and I think going to high Z allows us to check individual components in a very nice way.

And also I should mention that there has been a recent calculation by Desidero and Johnson of Notre Dame to recalculate the Lamb shift for the K-edge in mercury; they give a new result which agrees very well with experiment. In addition, G. Erickson is in the process of working out formulas which interpolate between high Z and low Z. So I think we will have plenty to check in this area.

J. E. Leiss: We have embarked on an experiment to try to—well, first test the suggestion of a proton halo, but in particular to measure the proton radius which your early paper on this subject pointed out is unknown by more than was previously thought. Could you give us an estimate of about how well we must measure the rms radius to be significant?

S. J. Brodsky: I think aside from the atomic physics Lamb shift test, it's very important to have the charge radius of the proton pinned down as hard as possible just from the point of view of the form factors of the proton. You know the neutron form factor is pinned down at $q^2=0$ by the accurate value for this slope. It's essential to have the value of the proton form factor pinned down by an accurate value at low q^2. This will greatly improve the fitting of the form factors. I think it would be nice to have an absolute, reliable number at 1 per cent accuracy on the charge radius of the proton. There are fits that claim to be that accurate but I'm not sure if they are believable. It is also critical to measure the elastic form factor of the deuteron for the D Lamb shift.

J. E. Leiss: The proton is easier.

S. J. Brodsky: Is it? Okay.

J. E. Leiss: Once you know the proton form factor the proton-deuteron ratio experiment is relatively easy.

S. J. Brodsky: Also, there are inelastic contributions that come into the deuteron experiment.

Theory of the Positronium Hyperfine Structure*

Thomas Fulton, David A. Owen, and Wayne W. Repko*****

Department of Physics, The Johns Hopkins University, Baltimore, Md. 21218

We present a calculation of the $m\alpha^6 \ln\alpha^{-1}$ contributions to the triplet-singlet splitting of the positronium ground state (the positronium "hyperfine" structure). Our result,

$$\Delta\nu(\alpha^6 \ln\alpha^{-1}) = \tfrac{3}{4}\alpha R_\infty \ln\alpha^{-1} = 34 \text{ MHz}.$$

places the theoretical value for the transition frequency ν approximately one standard deviation above the experimental value. The calculation is performed using a perturbation theory based on the Bethe-Salpeter equation, and a wave function obtained by a single iteration from the non-relativistic Pauli wave function.

Key words: Bethe-Salpeter equation; positronium hyperfine structure.

1. Introduction

Potentially, the splitting between the triplet and singlet levels of the positronium ground state affords a very accurate test of quantum electrodynamics. This follows since the bound electron-positron system, perhaps more than any other, presents a very nearly pure quantum electrodynamical problem, which to our order of accuracy is free from contamination by hadronic effects or unusual leptonic effects. Therefore, a favorable agreement between the experimental and theoretical determinations of the triplet-singlet splitting is necessary in any systematic check of the predictive power of quantum electrodynamics. Furthermore, positronium is the only experimentally accessible bound system which can and must be described by using the relativistic, two-body equation for interacting fermions. Agreement between theory and experiment is, therefore, also a check of the treatment of bound states in field theory.

Motivated by these considerations, we have calculated the corrections to the triplet-singlet transition frequency ν of order $\alpha^2 \ln\alpha^{-1}$. This calculation is necessary since the accuracy of the previous theoretical value for this separation [1, 2, 3] has been exceeded by that of the most recent experimental value [4]. Hence, a meaningful comparison between theory and experiment could not be made until higher order corrections were taken into account.

Our result for the $\alpha^2 \ln\alpha^{-1}$ correction to the transition frequency ν is

$$\Delta\nu(\alpha^6 \ln\alpha^{-1}) = \tfrac{3}{8} m\alpha^6 \ln\alpha^{-1}, \qquad (1)$$

where $\hbar = c = 1$. Combining this contribution with the order α correction of Karplus and Klein [3], the total triplet-singlet transition frequency becomes

$$\nu = \alpha^2 R y_\infty [\tfrac{7}{6} - (\tfrac{16}{9} + \ln 2)(\alpha/\pi) + \tfrac{3}{4}\alpha^2 \ln\alpha^{-1} + 0(\alpha^2)]. \qquad (2)$$

The contributions of order $m\alpha^6 \ln\alpha^{-1}$ represent recoil corrections arising from low momentum components of the wave function associated with the Bethe–Salpeter (B–S) equation [5] for positronium. The perturbation techniques used by Karplus and Klein [3] and Fulton and Martin [6] prove sufficiently accurate to this order, and we obtain the necessary wave function from the B–S equation, with only the Coulomb interaction, by a single iteration from the non-relativistic Pauli wave function. This wave function is expressed to a higher order in the wave function momentum than was previously necessary. We must also keep other momentum contributions to a higher order. Wave function retardation effects are only important in the one photon exchange process, while wave function pair effects can be neglected to our order. It is, however, essential that these two effects are taken into account in the interaction kernels.

In the following, we denote the space-time coordinates by $x_\mu = (\mathbf{x}, ix_0)$ and take the γ_μ to be hermitian matrices satisfying $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$.

2. Perturbation Theory

We begin by presenting the basic results of the perturbation theory as developed by Fulton and Karplus [7]. The essential idea is to separate the instantaneous Coulomb interaction, which accounts for the major part of the binding, from the total

* Supported in part by the National Science Foundation.

** Present address: Department of Physics, University of Surrey, Guildford, Surrey, England.

*** Present address: Department of Physics, Michigan State University, East Lansing, Michigan 48823.

interaction. With the aid of an approximate solution to the B–S equation containing the Coulomb interaction, we can then treat the contributions of the remaining time-dependent interaction terms by perturbation theory.

After mass and charge renormalization, the integral form of the B–S equation for the wave function $\psi(x_1, x_2)$ of an interacting two fermion system is:

$$\psi(x_1, x_2) = S_F^{(1)}(x_1 - x_1') S_F^{(2)}(x_2 - x_2') \times I(x_1', x_2'; x_3, x_4)\psi(x_3, x_4), \quad (3)$$

where $S_F(x)$ is the propagator for a spin $\frac{1}{2}$ particle, and repeated four-vector arguments are integrated over. The (irreducible) interaction kernel $I(x_1, x_2; x_3, x_4)$ is defined by the expansion of the electron-positron Green's function $G(x_1, x_2; x_3, x_4)$ in a perturbation series.

For purposes of calculation, it is convenient to introduce center-of-mass and relative coordinates X and x defined by

$$X = \tfrac{1}{2}(x_1 + x_2), \qquad x = x_1 - x_2. \quad (4)$$

Then, noting that $S_F^{(1)}(x_1 - x_1') S_F^{(2)}(x_2 - x_2')$ and $I(x_1, x_2; x_1' x_2')$ are functions of x, x' and $X - X'$, we can eliminate the center-of-mass coordinates and write eq (3) in the form

$$\varphi_K(x) = G_K(x, y) I_K(y, y') \varphi_k(y'), \quad (5)$$

where K denotes the center-of-mass energy-momentum four-vector. We work in the rest system of the center-of-mass so that $K_\mu = (0, iK_0)$.

The perturbation theory is developed by writing eq (5) in the form

$$\varphi_K(x) = G_k(x, y)[I^c(y, y') + I_k'(y, y')]\varphi_K(y'), \quad (6)$$

where

$$I^c(y, y') = -[i\alpha\delta(y - y')\delta(y_0)/r]\gamma_4^{(1)}\gamma_4^{(2)} \quad (7)$$

is the instantaneous Coulomb interaction kernel and $I_K'(y, y')$ denotes the remaining part of the interaction kernel, which is treated as a perturbation. In the present calculation, $I_K'(x, x')$ consists of all processes which involve the exchange of one or two photons.

The unperturbed wave function $\varphi_{Kc}(x)$ and corresponding energy K_0^c are obtained by solving eq (6) with $I_K'(y, y')$ set equal to zero. It can then be shown [7] that the energy shift $\Delta E = K_0 - K_0^c$, to our order of accuracy, is expressible as

$$\Delta E = -i\bar{\varphi}_{Kc}(x)[I_K'(x, x') + I_K'(x, y) \times G^c(y, y') I_K'(y', x')]\varphi_{Kc}(x'), \quad (8)$$

where $G^c(y, y')$ denotes the Green's function for the B–S equation containing the Coulomb interaction. In order to make use of eq (8), we need an approximate form for the Coulomb wave function $\varphi_{Kc}(x)$. This may be obtained by using eq (6), with only the Coulomb interaction, as the basis of an iteration procedure [8] beginning with the non-relativistic Pauli wave function. After one iteration, $\varphi_{Kc}(x)$ can be expressed in the form

$$\varphi_{Kc}(x) = \varphi_c(x) + \delta\varphi_c(x) + \delta\varphi_c'(x), \quad (9)$$

where $\varphi_c(x)$ is the wave function used to obtain the α^5 contributions [3], while $\delta\varphi_c(x)$ and $\delta\varphi_c'(x)$ represent higher order correction terms. Explicit expressions for these wave functions are given in the appendix.

3. Single Transverse Photon Contributions

3.1. Pure One Photon Exchange

The simplest correction term which produces an $\alpha^6 \ln\alpha^{-1}$ contribution to the triplet-singlet splitting is the single transverse photon exchange kernel $I_B(x, x')$. This (Breit) interaction kernel has the form

$$I_B(x, x') = \frac{ie^2\delta(x - x')\gamma_4^{(1)}\gamma_4^{(2)}}{(2\pi)^4} \times \int d^4k \, \frac{(\alpha^{(1)} \cdot \alpha^{(2)} - \alpha^{(1)} \cdot \hat{\mathbf{k}}\alpha^{(2)} \cdot \hat{\mathbf{k}})}{k^2} e^{ik \cdot x}, \quad (10)$$

where $\hat{\mathbf{k}} = \mathbf{k}/|\mathbf{k}|$. Denoting the single transverse photon exchange contribution to the triplet-singlet splitting by ΔE_1, we may use eq (8) and the expression for $\varphi_{Kc}(x)$ to obtain

$$\Delta E_1 = -i\bar{\varphi}_c(x) I_B(x, x')\varphi_c(x') - i\delta\bar{\varphi}_c(x) I_B(x, x')\varphi_c(x') - i\bar{\varphi}_c(x) I_B(x, x')\delta\varphi_c(x'), \quad (11)$$

where we have omitted $\delta\varphi_c'$, which can easily be shown to yield no $\alpha^6 \ln\alpha^{-1}$ contributions. The first term in eq (11) contains a spin dependent $\alpha^6 \ln\alpha^{-1}$ contribution given by

$$-i\bar{\varphi}_c(x) I_B(x, x')\varphi_c(x') = \tfrac{1}{12} m\alpha^6 \ln\alpha^{-1} \langle \boldsymbol{\sigma}^{(1)} \cdot \boldsymbol{\sigma}^{(2)} \rangle. \quad (12)$$

Remarkably, the remaining terms in eq (11) do not contribute due to a relative minus sign in the time dependences of $\varphi_c(x)$ and $\delta\varphi_c(x)$ (see appendix eqs (25) and (26)).

By calculating the spin matrix element in eq (12), we find

$$\Delta E_1 = \tfrac{1}{3} m\alpha^6 \ln\alpha^{-1}. \quad (13)$$

3.2. Coulomb-Transverse Exchange

In computing the effects of a single transverse photon exchange, it is important to remember that, because of retardation, the exchange of a single

transverse photon can be accompanied by the exchange of instantaneous Coulomb photons. This is evident in the α^5 calculation, where an important cancellation occurs between the one transverse photon exchange diagram and the diagram containing one transverse photon and one Coulomb photon [3]. A similar cancellation will be seen to occur in the present calculation.

If we denote the Coulomb transverse kernel by $I_{CT}(x, x')$, and the corresponding energy shift by ΔE_{CT}, then the $\alpha^6 \ln\alpha^{-1}$ spin dependent part of ΔE_{CT}, is given by

$$\Delta E_{CT} = -i \int d^4x d^4x' [\bar{\varphi}_P(0) I_{CT}(x, x') \varphi_c(x') + \bar{\varphi}_c(x) I_{CT}(x, x') \varphi_P(0)], \quad (14)$$

where $\varphi_P(0)$ is the amplitude of the non-relativistic wave function at the origin. The evaluation of eq (14) yields

$$\Delta E_{CT} = -\tfrac{1}{3} m\alpha^6 \ln\alpha^{-1}. \quad (15)$$

4. Double Transverse Photon Contributions

Recalling eq (8), we see that contributions to ΔE involving two transverse photons can arise in two ways. First, there is the contribution from the irreducible crossed two photon kernel $\tilde{I}_K(x, x')$. Next, there is the effective uncrossed two photon kernel $I_K(x, x')$, corresponding to the iteration of the single transverse exchange kernel $I_B(x, x')$. Explicitly, $I_K(x, x')$ is

$$I_K(x, x') = I_B(x, y) G_K(y, y') I_B(y', x'). \quad (16)$$

The evaluation of the contributions of $I_K(x, x')$ and $\tilde{I}_K(x, x')$ can be simplified by noting that the addition of another photon introduces a factor α into the expression for the energy shift. This allows us to replace one of the wave functions in eq (8) by $\varphi_P(0)$, thereby obtaining

$$\Delta E_2 = -i \int d^4x d^4x' \{\bar{\varphi}_P(0)[I_K(x, x') + \tilde{I}_K(x, x')] \times \varphi_c(x') + \bar{\varphi}_c(x)[I_K(x, x') + \tilde{I}_K(x, x')]\varphi_P(0)\}, \quad (17)$$

where ΔE_2 denotes the energy shift associated with the two transverse photon exchange. Further, it is possible to neglect the time dependence of $\varphi_c(x)$. With these simplifications, eq (17) can be evaluated to give

$$\Delta E_2 = \tfrac{3}{8} m\alpha^6 \ln\alpha^{-1}. \quad (18)$$

5. Noncontributing Diagrams

In addition to the contributions already discussed, there exist numerous other diagrams which are known to give contributions in order α^5. These consist of radiative corrections to the single photon exchange, single and double photon annihilation terms and the crossed-Coulomb terms. We have carried out a systematic check of these diagrams and find that they do not contribute [9] to order $\alpha^6 \ln\alpha^{-1}$. The essential reason for the absence of $\alpha^6 \ln\alpha^{-1}$ corrections in these diagrams is that the integrals involved in their evaluation are not sufficiently singular in the region of low wave function momentum. This lack of singularity can be attributed to the fact that all these processes are relatively high energy effects, which are only important for large values of the wave function momentum. The details of these calculations can be found in ref. [9].

Since diagrams involving three photons are of order α^6, we conclude that the $\alpha^6 \ln\alpha^{-1}$ contributions are limited to those given above.

6. Summary and Discussion

Combining eqs (13), (15) and (18), we obtain the total triplet-singlet energy separation

$$\Delta E = \Delta E_1 + \Delta E_{CT} + \Delta E_2 = \tfrac{3}{8} m\alpha^6 \ln\alpha^{-1}. \quad (19)$$

On taking the value of α^{-1} to be [10]

$$\alpha^{-1} = 137.03608, \quad (20)$$

eq (19) leads to a frequency shift $\Delta\nu$ given by

$$\Delta\nu(\alpha^6 \ln\alpha^{-1}) = 34 \text{ MHz}. \quad (21)$$

When this correction is added to the previous result of Karplus and Klein [3], we obtain

$$\nu\text{th (Karplus and Klein)} = 2.03381 \times 10^5 \text{ MHz}$$

$$\nu\text{th (present paper)} = 2.03415 \times 10^5 \text{ MHz}, \quad (22)$$

which is to be compared with the experimental value

$$\nu^{\text{exp}}(1970) = 2.03403(12) \times 10^5 \text{ MHz}. \quad (23)$$

One observes that the contributions we obtain raise the theoretical value from two standard deviations below to one standard deviation above the experimental value. Considering the estimated $m\alpha^6$ contributions given in table 1, the agreement between theory and experiment seems quite reasonable [11].

TABLE 1. *Theoretical contributions to positronium frequencies*

	Order	$\Delta\nu$ 10^5 MHz	Percentage of contribution of last order	
			Actual	Expected
Schroedinger Level (ionization frequency)	α^2	16449.2	—	—
Triplet-singlet ground state splitting	α^4	2.04386	0.012	0.005
	α^5	−0.01005	0.49	0.73
	$\alpha^6 \ln\alpha^{-1}$	0.00034	3.4	3.6
	α^6 (est)	±0.00007	—	20

7. Appendix

Introducing the Fourier transform of $\varphi_{Kc}(x)$ by

$$\varphi_{Kc}(x) = (2\pi)^{-3/2} \int d^3p \varphi_{Kc}(\mathbf{p}, t) \exp(i\mathbf{p}\cdot\mathbf{x}) \quad (24)$$

the various terms in eq (9) are given by

$$\varphi_c(\mathbf{p}, t) = 2\alpha(2\pi)^{-1/2} \frac{m}{E}\left(1+\frac{\alpha^{(1)}\cdot\mathbf{p}}{2m}\right)\left(1-\frac{\alpha^{(2)}\cdot\mathbf{p}}{2m}\right)$$

$$\times[mf_+(t)+Ef_-(t)]\frac{\varphi_P(0)}{(p^2+\gamma^2)^2} \quad (25)$$

$$\delta\varphi_c(\mathbf{p}, t) = \tfrac{1}{2}\alpha(2\pi)^{-1/2}p^2\frac{[mf_+(t)-Ef_-(t)]}{mE}\frac{\varphi_P(0)}{(p^2+\gamma^2)^2} \quad (26)$$

$$\delta\varphi_c'(\mathbf{p}, t) = -\tfrac{1}{2}\alpha(2\pi)^{-1/2}p^2(\alpha^{(1)}+\alpha^{(2)}) \cdot\mathbf{p}(f_-(t)/mE)[\varphi_P(0)/(p^2+\gamma^2)^2], \quad (27)$$

where

$$E=(p^2+m^2)^{1/2}, \qquad \gamma=\tfrac{1}{2}m\alpha$$

and

$$f_\pm(t) = \tfrac{1}{2}\{\exp[-i(E-m)\,|\,t\,|] \pm \exp[-i(E+m)\,|\,t\,|]\}. \quad (28)$$

8. References

[1] Pirenne, J., Arch. Sci. Phys. Nat. **28**, 233 (1946), and **29**, 121, 207, 265 (1947).
[2] Ferrell, R. A., Phys. Rev. **84**, 858 (1951).
[3] Karplus, R., and Klein, A., Phys. Rev. **87**, 848 (1952).
[4] Deutsch, M., and Brown, S. C., Phys. Rev. **85**, 1047 (1952); Weinstein, R., Deutsch, M., and Brown, S. C., Phys. Rev. **94**, 758 (1954) and **98**, 223 (1955); Hughes, V. W., Marder, S., and Wu, C. S., Phys. Rev. **106**, 934 (1957); Theriot, E. D., Jr., Beers, R. H., and Hughes, V. W., Phys. Rev. Lett. **18**, 767 (1967); Theriot, E. D., Jr., Beers, R. H., Hughes, V. W., and Ziock, K. O. H., Phys. Rev. A (1970) (to be published).
[5] Salpeter, E. E., and Bethe, H. A., Phys. Rev. **84**, 1232 (1951).
[6] Fulton, T., and Martin, P. C., Phys. Rev. **93**, 903 and **95**, 811 (1954).
[7] Fulton, T., and Karplus, R., Phys. Rev. **93**, 1109 (1954).
[8] Salpeter, E. E., Phys. Rev. **87**, 328 (1952).
[9] Owen, D. A., Ph.D. Thesis, The Johns Hopkins University (unpublished).
[10] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[11] Table 1 was computed using the method of Fulton, T., Owen, D. A., and Repko, W. W., Phys. Rev. Lett. **24**, 1035 (1970). Equation (21) of Section 6 above includes a retardation effect which is not contained in this reference.

DISCUSSION

A. Ruark: I have two points which might be either shocking or stimulating to the experimentalists and to the Hopkins group. First of all, in many of the experiments to check up these very interesting calculations, the atoms are always swimming in a nice sea of blackbody radiation, aren't they? Now, what will we do about that? The second point is perhaps more of a fundamental shocker. Suppose the electron and the positron are not truly identical. And there is a certain small minority of people who worry about such things.

D. A. Owen: Well, I think that if some of these things are true, then we will expect to find discrepancies between our calculations and the experimental results.

W. E. Lamb, Jr.: I think I might make a partial answer to one of Dr. Ruark's points. The presence of some radiation, like blackbody radiation, can be allowed for, theoretically, in some of the experiments. It doesn't seem to be a very big correction, but it should be looked at again every time a new experiment is done.

M. Danos: Is it true that the course of the corrections resulting from intermediate states containing mesons have not been taken into account? And if so, has the correction taking those into account about the same magnitude as those in the alpha 6 correction?

D. A. Owen: Oh. No, that doesn't contribute here. We have looked at the vacuum polarization. In other words, the self-energy of the photon, and have found no $\alpha^6 \ln\alpha^{-1}$ contribution.

Precision Measurement of the Fine Structure Interval of the Ground State of Positronium*

E. R. Carlson, V. W. Hughes, and E. D. Theriot, Jr.[1]

Gibbs Laboratory, Yale University, New Haven, Connecticut 06520

An experiment is in progress to improve the accuracy of the determination of the fine structure interval of the ground state of positronium to about 10 ppm from its present level of 59 ppm. The present status of this experiment is discussed. An accuracy of 10 ppm will provide a sensitive test of the $\alpha^4 \mathrm{Ry}\ \ln\alpha^{-1}$ term in the theoretical fine structure interval and will necessitate a calculation of the terms of order $\alpha^4 Ry$.

Key words: Positronium fine structure.

1. Introduction

Positronium, the bound state of an electron and a positron, is an ideal system for a test of quantum electrodynamics because only a lepton and its antiparticle are present. The study of positronium provides the principal test of the Bethe-Salpeter equation, which describes the bound state quantum electrodynamic two-body system. The fine structure interval of the ground state, $\Delta\nu$, is the quantity that has been measured with precision. The theory of $\Delta\nu$ has been calculated to the order $\alpha^4 \ln\alpha$ Ry.

The current experimental value for the fine structure interval of the ground state of positronium, $\Delta\nu$, has been obtained by measuring the microwave transition frequency between the $M=0$ and $M=\pm 1$ Zeeman levels of orthopositronium in a static magnetic field of about 7900 G [1]. The technique used is basically similar to that of previous measurements [2, 3].

2. Theory of the Experiment [4]

Figure 1 shows the Zeeman energy levels of the ground state of positronium. The quantity to be determined is $\Delta\nu$, the separation between the $^3S_1$ state, or orthopositronium, and the $^1S_0$ state, or parapositronium.

Because positronium is a particle-antiparticle system, it annihilates into gamma rays. The annihilation properties of the ortho- and para-states are different: orthopositronium has a mean life of 1.4×10^{-7} s and decays into three coplanar gamma rays with total energy 1.02 MeV; parapositronium has a mean life of 1.25×10^{-10} s and decays into two gamma rays 180° apart in direction, each with energy 511 keV.

When a static magnetic field is applied to positronium, the $M=\pm 1$ triplet levels are not affected (because the atom has no net magnetic moment in this state) but the $M=0$ triplet and singlet wavefunctions are mixed and the energy levels are shifted. The admixture of singlet wavefunction causes the

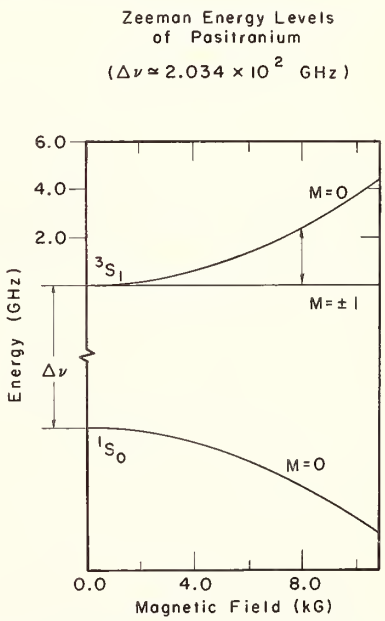


FIGURE 1. *Zeeman energy levels of positronium in the ground state.*

$M=0$ triplet state to decay predominantly into two gamma rays and shortens the mean life to about 10^{-8} s at a field of 7900 G.

Measurement of $\Delta\nu$ by inducing the direct transition between the $^3S_1$ and $^1S_0$ states is difficult because a high power microwave source at 200 GHz is

* This research has been supported in part by the National Science Foundation. (Grant GP-23722.)

[1] Present address: National Accelerator Laboratory, Batavia, Illinois 60510.

required. The fine structure interval can be expressed, however, in terms of the Zeeman level separation, the magnetic field, and fundamental constants, using the Breit-Rabi formula:

$$f_{01} = (\Delta\nu/2)[(1+x^2)^{1/2} - 1] \qquad (1)$$

where

$$x = 2\mu_0 gH/h\Delta\nu \qquad (2)$$

and f_{01} is the Zeeman transition frequency in orthopositronium, μ_0 is the Bohr magneton, g is the electron g-value, H is the static magnetic field and h is Planck's constant. Measurement of the Zeeman level separation in a known magnetic field allows a determination of $\Delta\nu$. At 7900 G the Zeeman frequency is about 2.4 GHz for which convenient sources of power are available. There is no fundamental disadvantage to studying the Zeeman transition as compared to the direct $\Delta\nu$ transition, because the ratio of the natural radiative width to the resonance frequency is the same for both transitions.

The experimental technique, then, is to form positronium in a magnetic field, induce Zeeman transitions from the $M = \pm 1$ to the $M = 0$ states of orthopositronium with a microwave magnetic field, and detect the transitions through the increase in two quantum annihilations. It is also possible to induce transitions from the $M = 0$ to $M = \pm 1$ states, but the probability is much smaller due to the short lifetime of the $M = 0$ state.

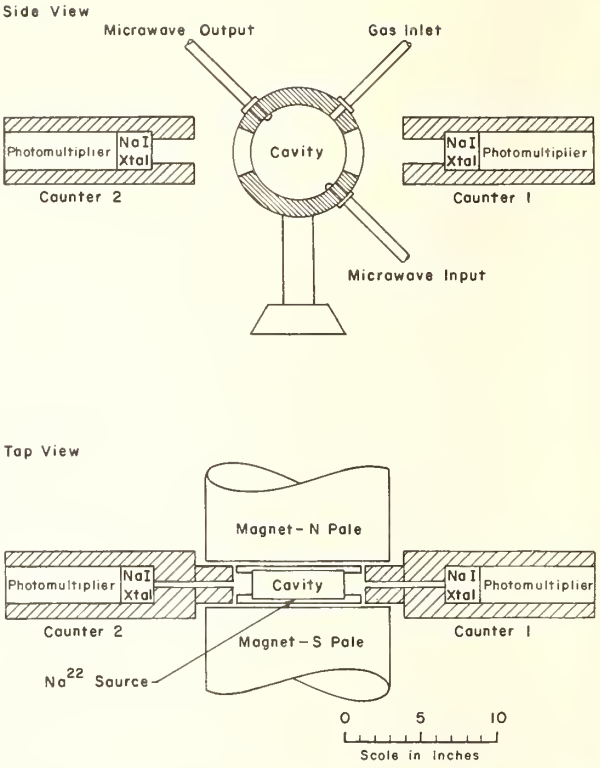


FIGURE 2. *Schematic diagram of experimental apparatus for positronium fine structure measurement.*

3. Experimental Apparatus

Figure 2 shows the experimental apparatus. The static magnetic field is produced by an electromagnet. It is important that the field be homogeneous because the region in which positronium annihilates is not highly localized. The field is stabilized by an nmr system whose frequency is derived from a crystal-controlled oscillator.

The microwave field is produced inside a cavity which is positioned between the magnet pole pieces. The cavity is operated in the TM_{110} mode which produces a strong field with approximately linear polarization and constant amplitude in the region of positronium formation. The microwave frequency is synthesized by multiplication of the output from a crystal-controlled oscillator, and power is provided by a klystron amplifier which can deliver 500 W into the cavity through the input coupler. The output coupler samples the field in the cavity and is the input to a power leveling servo-system. High power and a high-Q cavity are necessary to achieve a reasonable transition rate in a short-lived system with a magnetic dipole transition.

The positron source is 10 mCi of sodium-22 placed in a hole in the center of one cavity wall. Positrons emitted into the cavity are confined by the magnetic field to a cylindrical volume with approximately 1 in diam centered on the cavity axis. The positrons from ^{22}Na have a maximum energy of 542 keV, and a most probable energy of 120 keV. These energetic positrons are slowed down by inelastic collisions with an inert gas introduced at the gas inlet, typically ultra high purity argon at pressures up to five atmos. The slow positrons capture electrons from the gas atoms and form positronium. Ortho- and para-positronium are formed in the ratio of three to one except for a small population difference due to polarization of the positrons from the source.

The annihilation gamma rays are detected by sodium iodide crystals coupled to photomultiplier tubes. The detectors are shielded from the magnetic field and from scattered gamma rays. In front of each detector, lead blocks are arranged to form a collimating slit so that the gamma rays from annihilations in the cavity walls cannot be observed directly. Two detectors are operated as a coincidence pair using fast logic, and are located to ensure observation of the two quantum annihilation of positronium. In order to be accepted by the system, an event must consist of two gamma rays each having energy 511 keV, 180° apart in space, and coincident in time. To increase sensitivity to changes in the two quantum annihilation rate without becoming susceptible to drifts in photomultiplier gain or logic thresholds, the pulse from one detector is stored in a multichannel pulse height analyzer when an event is accepted. A computer program fits a curve to the photopeak in the spectrum and performs an integration over specified limits to yield the two quantum counting rate.

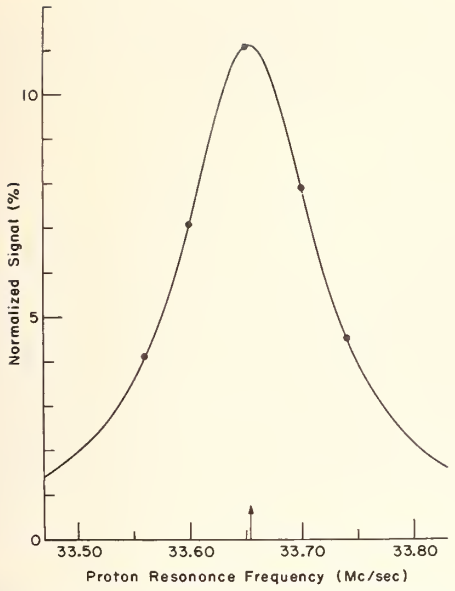


FIGURE 3. *A typical resonance line for the Zeeman transition in orthopositronium.*
Data were taken with argon pressure = 45 lb/in², microwave frequency = 2385.045 MHz, and microwave power = 100 W.

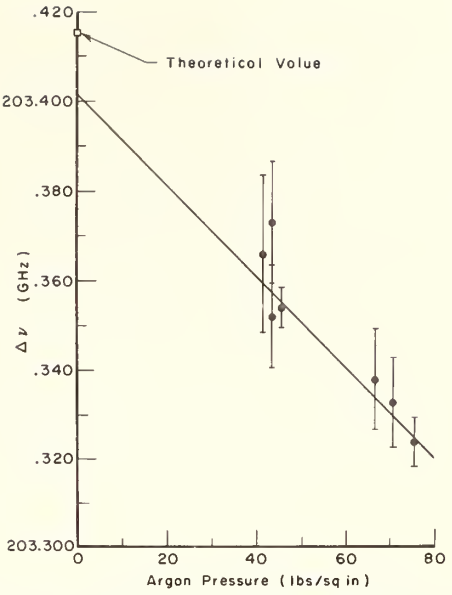


FIGURE 4. *Measured values of Δν versus argon pressure.*
The solid curve is a straight line fit to the data.

Figure 3 shows a typical fitted resonance line shape. The abscissa is static magnetic field measured in units of the proton resonance frequency of the nmr system. The ordinate is normalized signal which is related to the probability that a three-quantum annihilation is converted to a two-quantum annihilation by the microwave field. The dots are experimental points with statistical errors approximately the size of the dots. The curve is a best fit of the theoretical line shape, which involves the three orthopositronium levels $M = 0, \pm 1$. The linewidth is about three parts in 10^3 and is due to the natural width associated with annihilation and to microwave power broadening.

4. Summary of Results

Figure 4 shows a plot of $\Delta\nu$ versus pressure of the argon gas in the experiment of Theriot, Beers, and Hughes [1]. A straight line is fitted to the experimental points, and $\Delta\nu$ for free positronium is taken as the extrapolated value at zero pressure. The current experimental value of $\Delta\nu$ is

$$\Delta\nu_{\text{expt}} = 203.403 \pm 0.012 \text{ GHz [1]}. \quad (3)$$

The 60 ppm error quoted is one standard deviation and is due to statistical counting errors and magnetic field inhomogeneity.

The theoretical value [5] is

$$\Delta\nu_{\text{theor}} = \alpha^2 Ry_\infty \left[\tfrac{7}{6} - (\alpha/\pi)\left(\tfrac{16}{9} + \ln 2\right) + \tfrac{3}{4}\alpha^2 \ln\alpha^{-1}\right]$$
$$= 203.415 \text{ GHz} \quad (4)$$

using

$$\alpha^{-1} = 137.03608 \quad (5)$$

The values of $\Delta\nu_{\text{expt}}$ and $\Delta\nu_{\text{theor}}$ agree to within one standard deviation of the experimental error.

5. Further Development of the Measurement

An experiment is in progress at Yale University to improve the accuracy of $\Delta\nu_{\text{expt}}$ to about 10 ppm.

The most serious limitation on the accuracy of the previous measurements is counting statistics, so improvements have been made in the experiment to increase the data acquisition rate. Four pairs of coincidence detectors are used, and the solid angle per detector has been increased by using $2\frac{1}{2}$ in $\times 2\frac{1}{2}$ in NaI crystals and by changing the collimator design. The crystals are coupled to photomultiplier tubes through acrylic light pipes so that the photomultipliers and their magnetic shields are removed from the immediate vicinity of the magnet gap, thus preventing gain shifts due to magnetic field changes and deterioration of the field homogeneity. A strong positron source will also increase the counting rate. Coincidence resolution of 200 ns is achieved using a slow-fast coincidence system which reduces the rate of accidental coincidences. The coincidence spectra and experimental parameters are recorded by an automatic data acquisition system.

Improvement of the accuracy in the determination of $\Delta\nu$ also requires a more homogeneous magnetic field. The current experiment uses a 15-in Varian magnet with a 4-in air gap. It has been shimmed in the laboratory to provide a field with a total variation of 8 ppm at 7.8 kG over the 1 $(\text{in})^3$ volume in which positronium is observed. Field stability of a few parts in 10^7 has been achieved using an nmr system. In addition, care is taken to minimize tem-

perature fluctuations of the magnet which could result in changes in field distribution.

The wider air gap of the new magnet has permitted an increase in the volume in which positronium is observed by a factor of about two. This increases the counting rate in comparison with the previous experiment and also provides a useful signal at lower stopping gas pressure. Measurement of $\Delta\nu$ at lower pressure reduces the error resulting from the extrapolation required by the fine structure pressure shift. In order to minimize possible systematic errors from the fine structure shift, data will be taken in several gases and for a range of pressures.

The 10 ppm accuracy hoped for in the new measurement will be limited by statistics and magnetic field inhomogeneity. It will require determination of the center of the resonance line to about one part in 750 of its natural width.

An accuracy of 10 ppm in the measurement of $\Delta\nu$ will provide a sensitive test of the $\alpha^4 \ln\alpha$ Ry term in the theory and will necessitate a calculation of the terms of order α^4 Ry.

6. References

[1] Theriot, E. D., Jr., Beers, R. H., and Hughes, V. W., Phys. Rev. Letters **18**, 767 (1967); Theriot, E. D., Jr., Beers, R. H., Hughes, V. W., and Ziock, K. O. H., Phys. Rev. A **2**, 707 (1970).

[2] Deutsch, M., and Brown, S. C., Phys. Rev. **85**, 1047 (1952); Weinstein, R., Deutsch, M., and Brown, S., Phys. Rev. **94**, 758 (1954); **98**, 223 (1955).

[3] Hughes, V. W., Marder, S., and Wu, C. S., Phys. Rev. **106**, 934 (1957).

[4] Deutsch, M., Progr. Nucl. Phys. **3**, 131 (1953).

[5] Fulton, T., Owen, D. A., and Repko, W. W., Phys. Rev. Letters **24**, 1035 (1970); see also these Proceedings.

DISCUSSION

G. W. Series: I wonder if you have to apply any corrections of the Bloch-Siegert type for the other component of the oscillating field, or alternatively, corrections due to the presence of that other energy level. This might be important if the fields weren't quite properly alined. Are these important corrections?

E. R. Carlson: The Bloch-Siegert correction, no, not to the results so far. I'm not sure what you mean by the "other energy level."

G. W. Series: You have transitions from 0 to plus or minus 1.

E. R. Carlson: Right.

G. W. Series: But the 0 to 0 would go if you had the fields parallel, wouldn't it? And you are not in resonance, so there would be a slight correction due to the existence of that state.

E. R. Carlson: We are a factor of 100 off resonance, and I believe that correction is too small to be important in this value.

S. J. Brodsky: I sort of doubt that the Breit-Rabi formula for the positronium Zeeman structure could be reliable to a few parts per million. I think that there is considerable work that has to be done starting from the Bethe-Salpeter formalism for the two-body system in an external field. I know in the case of hydrogen the mass ratio eliminates such problems, but for positronium I'm quite suspicious that you would find corrections at a few parts per million.

W. E. Lamb, Jr.: Very often when you try to find errors in the Breit-Rabi formula and you have allowed for changes in ΔE's and g value, you find you still have the right Breit-Rabi formula with the changed zero field splittings and g values but that may not be a universal law.

Fine and Hyperfine Structures of Atomic Hydrogen

Norman F. Ramsey

Department of Physics, Harvard University, Cambridge, Mass. 02138

The fine and hyperfine structures of the atomic hydrogen spectrum have been particularly fruitful sources of new developments in physics and the precise measurements of these structures have provided accurate checks on quantum electrodynamics and atomic theory. During the past few years there have been new measurements and new theoretical calculations of the fine and hyperfine structures of atomic hydrogen. These new measurements and theories have removed most of the discrepancies that previously existed between experiment and theory and there is now remarkable agreement between experiment and theory.

Key words: Atomic hydrogen; atomic deuterium; hyperfine structure; fine structure; hydrogen maser; fine structure constants.

1. Introduction

Since atomic hydrogen is the simplest of all atoms to interpret theoretically, it is not surprising that studies of the fine and hyperfine structure of the atomic hydrogen spectra have been a fruitful source of new ideas in physics. The fine structure of atomic hydrogen, for example, through the observed Lamb shift [1] provided the experimental stimulation to the revolution in quantum electrodynamics [2] that began in the late 1940's. The hyperfine structure on the other hand provided the first evidence for the anomalous magnetic moment of the electron [3] and for further quantum electrodynamic corrections. Both the fine and the hyperfine spectral studies have yielded accurate values [4, 5] of the fine structure constant α.

The agreement between theory and experiment in both fine and hyperfine measurements has in many ways been remarkable. However, up to the current year there have been two nagging discrepancies: (a) the values of the fine structure constant α inferred from the fine structure experiments [6–16] has differed [4–5] from those calculated from hyperfine structure measurements [18–24] by an amount well beyond experimental and theoretical error [2, 4, 5] and (b) there have been small but significant discrepancies between theory and experiment on the Lamb shift [2, 4, 5]. Now for the first time there is remarkably better agreement on all these values. Four developments have contributed most to this improved situation: (a) Improved calculations by Iddings, Yennie and others showed the discrepancies could not be attributed to uncertainties in the proton structure correction to the hydrogen hyperfine structure, (b) the independent measurements [5, 25] of α by the Josephson effect with a full confirmation of the hyperfine value [4, 18–24] of α, (c) recent improved experimental measurements [6–16] which made the previous discrepancies between theory and experiment even more acute, and (d) the discovery by Applequist and Brodsky [2] of a 0.35 MHz error in previous theoretical calculations of the Lamb shift.

As a result of these recent developments, it is now possible to present the observed data for both the fine and hyperfine experiments on H, D, and T in the form of a comparison between experiment and theory with a single set of fundamental constants including a single α being used in the theoretical calculations and with excellent agreement between experiment and theory.

2. Fine Structure of Atomic Hydrogen

From the Dirac theory of the electron, it was anticipated that the $2^2P_{1/2}$ and the $2^2S_{1/2}$ states would be exactly degenerate. However, the experiment of Lamb and Retherford [1] showed conclusively that the $2^2S_{1/2}$ state was below that of the $2^2P_{1/2}$ by approximately one thousand MHz. With the inclusion of the Lamb shift, the energies of the $n=2$ states of atomic hydrogen and their variations with magnetic field are as given in figure 1, where for simplicity the small hyperfine structure splittings have been omitted. The spacings of these energy levels have been subsequently studied in both atomic hydrogen and atomic deuterium by a number of observers [6–17]. Most of the experiments involve the use of atomic beams and the use either of oscillatory magnetic fields to induce transition between the levels or of level crossing techniques in which the magnetic fields are such that transitions between the states occur due to their degeneracy. Since the different measurements are performed at different magnetic fields, the consistency of the final results not only gives values for the fine structure separations but also confirms the validity for the Zeeman effect as normally calculated for the fine structure and as given graphically in figure 1.

The results of the different measurements of the fine structure separations are given in table 1 along with theoretical calculations of the same quantities.

TABLE 1. *Comparison between experimental measurements and theoretical calculations of the fine structure of atomic hydrogen and its isotopes*

$\delta = {}^2P_{1/2} - {}^2S_{1/2}$ and $\Delta E = {}^2P_{3/2} - {}^2P_{1/2}$. The theoretical values are those of Applequist and Brodsky [2] with $\alpha^{-1} = 137.03608 \pm 0.0026$. The indicated errors are one standard deviation.

Atom	n	Interval	Experiment (MHz)	Theory (MHz)	Expt−Theory (MHz)	References
H	2	δ	1057.77±0.06	1057.91±0.06	−0.14±0.08	Triebwasser, et al. [6]
H	2	δ	1057.90±0.10	1057.91±0.06	−0.01±0.11	Robiscoe and Cosens [7]
H	2	ΔE–δ	9911.38±0.03	9911.12±0.07	+0.26±0.08	Kaufman, et al. [8]
H	2	ΔE–δ	9911.25±0.06	9911.12±0.07	+0.13±0.09	Shyn, et al [9]
H	2	ΔE–δ	9911.17±0.04	9911.12±0.07	+0.05±0.08	Vorburger and Cosens [10]
H	2	ΔE	10 969.6±0.7	10 969.03±0.04	+0.6±0.7	Wing [11]
H	2	ΔE	10 969.13±0.09	10 969.03±0.04	+0.10±0.11	Baird, et al. [12]
H	3	δ	313.6±5.7	314.7±0.1	−1.1±5.7	Kleinpoppen [13]
D	2	δ	1059.00±0.06	1059.17±0.07	−0.17±0.10	Triebwasser, et al. [6]
D	2	δ	1059.28±0.06	1059.17±0.07	+0.11±0.10	Cosens [14]
D	2	ΔE–δ	9912.59±0.10	9912.85±0.09	−0.26±0.15	Triebwasser, et al. [6]
D	3	δ	315.30±0.80	315.3	0.0±1.0	Wilcox and Lamb [16]
D	3	ΔE–δ	2934±10	2935	1±10	Lamb and Sanders [15]
D	3	ΔE	3250.7±2.0	3250.7	0±2	Wilcox and Lamb [16]
D	3	$P_{3/2}$–$D_{3/2}$	5±10	5.34	0±10	Wilcox and Lamb [16]
D	4	δ	133±10	133.1	0±10	Wilcox and Lamb [16]

The usual convention is followed of designating by δ the Lamb shift or energy separation ${}^2P_{1/2}$–${}^2S_{1/2}$ and by ΔE the energy separation ${}^2P_{3/2}$–${}^2P_{1/2}$. The theoretical values are those of Applequist and Brodsky [2] with $\alpha^{-1} = 137.03608 \pm 0.00026$. The indicated errors correspond to one standard deviation. The difference between the experimental and theoretical values are shown in the next to last column of table 1.

It can be seen that the agreement between theory and experiment is excellent. In those few cases where one experiment differs from theory by more than one or two standard deviations, there is another independent measurement of the same quantity which agrees with theory to much less than one standard deviation.

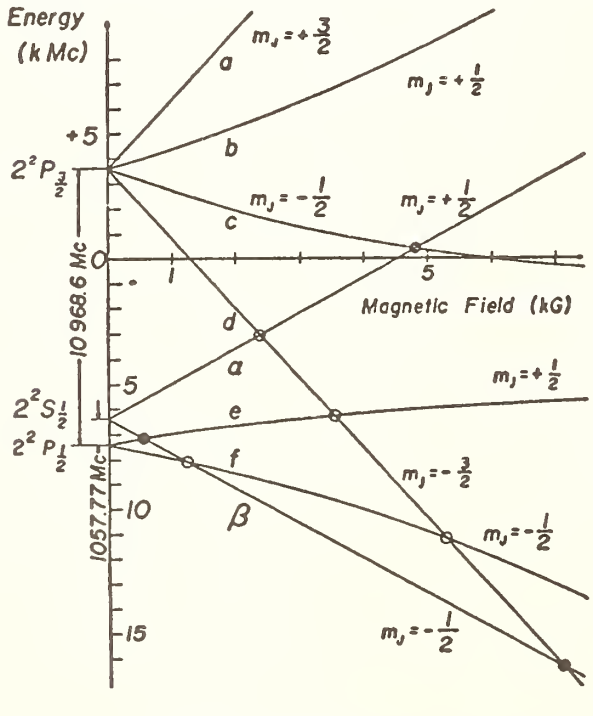


FIGURE 1. *Zeeman energy level diagram of the fine structure in the* n = 2 *state of hydrogen.*

3. Hyperfine Structure of Atomic Hydrogen

Various kinds of measurements have been made on the hyperfine structure of atomic hydrogen. Most of these have been for the $1^2S_{1/2}$ ground electronic state where the dependence of the energy levels on magnetic field are as shown in figure 2. The measurements have been by atomic beam methods [3, 27], by use of the hydrogen maser [18–24], or by optical pumping and spin exchange [28–30]. A comparison of the experimental results with those theoretically calculated are shown in table 2.

The most accurate values of the hyperfine separations $\Delta\nu$ are obtained from hydrogen maser studies and the values listed for H, D, and T are those obtained that way. These maser results are consistent with other less accurate measurements. The theoretical value for $\Delta\nu$ of H is calculated with the same value of α used in table 1. Originally the values of α obtained from fine and hyperfine measurements were different. However, the Josephson effect measurements [25] confirmed the hyperfine value for α and the recent theoretical correction of Applequist and Brodsky [2] to the theory of the Lamb shift brings the fine, hyperfine and Josephson effect measurements to consistency with the single value of α in tables 1 and 2.

With the consistency of all these measurements, the hyperfine experiments can be reinterpreted as a means to measure $\delta N^{(2)}$, the proton polarizability contribution [5] due to the various excited states and to the internal structure of the proton. As shown on the table, the theoretical and experimental values of $\delta N^{(2)}$ are consistent.

Except for one theoretical calculation for tritium,

TABLE 2. *Comparison between experimental measurements and theoretical calculations of the hyperfine structure of atomic hydrogen and its isotopes.*

The indicated errors correspond to one standard deviation. The theoretical values are from the indicated references with same α as table 1. δN^2 is the proton polarizability contribution due to excited states and internal structure. R is the ratio of hyperfine frequency in the 2S state to that in 1S.

Atom	Measurement	Experiment	Theory	Expt−Theory	References	
					Expt.	Theory
H	$\Delta\nu$	1,420,405,751.768±0.009 Hz	1,420,403,500±8000 Hz	2200±8000 Hz	18–22	5
H	$\delta N^{(2)}$	2.5±4.0 ppm	0±2 ppm	2.5±5.0 ppm	18–22	5
D	$\Delta\nu$	327,384,352.51±0.05 Hz	327,394,000±33,000 Hz	−10,000±33,000 Hz	31	32
T	$\Delta\nu$	1,516,701,470.7919±0.0071 Hz	1,516,695,000±8000 Hz	6400±8000 Hz	33	34, 51
			1,516,684,570±3000 Hz	17,000±3000 Hz	33	35, 52
H	$\delta\Delta\nu E$	$-(6.7\pm0.4)\times10^{-5}E^2$ Hz	$-7.63\times10^{-5}E^2$ Hz	$(0.9\pm0.6)E^2$ Hz	36	37, 38
H	$\delta\Delta\nu$H	Breit–Rabi formula confirmed to 5×10^{-9} oersted.				
H/D	g_J(H)/g_J(D)	$1+(9.4\pm1.4)\times10^{-9}$	$1+7.3\times10^{-9}$	$+(2.1\pm1.4)\times10^{-9}$	23	40, 41
H/D	g_J(H)/g_J(D)	$1+(7.2\pm3.0)\times10^{-9}$	$1+7.2\times10^{-9}$	$-(0.1\pm3.0)\times10^{-9}$	42	40, 41
H/T	g_J(H)/g_J(T)	Preliminary results intermediate between two D values.				
H	μ_e/μ_p	658.210705(6). Agrees with magnetic shielding theory.			43, 44	50, 53
H	$8R-1$	0.000034495(60)	0.00003445(2)	+0.000000045(60)	45, 54	46, 54
D	$8R-1$	0.0000342(6)	0.00003453(2)	−0.00000033(60)	49	48, 54

it is apparent that the deuterium and tritium values for $\Delta\nu$ are consistent with theory, but it should be emphasized that this consistency is not an accurate test of quantum electrodynamics since the accuracy of the theoretical calculations is limited by uncertainties in the corrections for the internal structure of the deuteron and triton.

The fifth row of table 2 is a comparison between the theoretical and experimental values for the dependence of $\Delta\nu$ upon an externally applied electric field. Again there is full consistency. Currently in our laboratory we are seeking to measure this quantity more accurately and to obtain a measurement of the frequency shift when the electric field is perpendicular rather than parallel to the magnetic field which defines the axis of quantization. There should be a small difference between these two values.

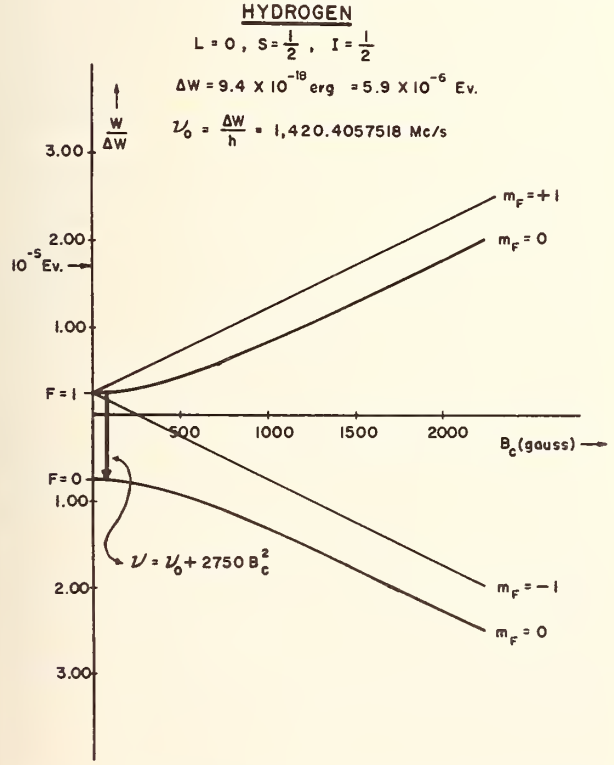


FIGURE 2. *Zeeman energy level diagram of the hyperfine structure for the $1^2S_{1/2}$ electronic ground state of atomic hydrogen.*

The sixth row pertains to a check of the theory of the dependence of the hyperfine separation upon an externally applied magnetic field. The theory of this dependence was developed many years ago by Breit and Rabi [50]. The most accurate comparison between theory and experiment was made at 8 Oe. The results experimentally confirm the Breit-Rabi formula for atomic hydrogen to an accuracy of 5×10^{-9}.

Recently there have been two independent measurements [23, 42] of the atomic magnetic moments of H and D and several new theoretical calculations [40, 41]. At present our value [23] differs from the latest theoretical prediction by 1.5 times the estimated error whereas that of Robinson and his associates [42] is in full agreement with the present theory. In our laboratory Larson is now in the process of studying the similar ratio for atomic tritium, and the preliminary value for tritium corresponds to a deuterium value intermediate between those of Robinson [42] and ourselves. Although there is no inconsistency at present between experiment and theory, we feel we have a major task ahead either to confirm or eliminate the small discrepancies which now exist.

In the third to the last row of table 1 we have listed the ratio between the magnetic moment of the electron and of the proton as measured in the hyperfine spectrum of a hydrogen maser operating at 3500 oersted. Although this value can not be accurately compared directly to theory due to uncertainties in the theoretical value of the proton magnetic moment, the result can be compared to experimental values

for the same ratio measured with protons in the H_2 molecule, where an important magnetic shielding correction must be made. The agreement of the results to the accuracy of the magnetic shielding correction both confirms the consistency of the atomic hydrogen calculations and provides the first absolute check of the validity of the author's theory of magnetic shielding [53].

The final two rows in table 2 pertain to the ratio R of the hyperfine frequency in the $2S$ state of hydrogen and deuterium to that in the $1S$. Again the agreement between theory and experiment is excellent.

4. Conclusions

It is apparent from the two preceding sections that there is now excellent agreement between theory and experiment for both quantum electrodynamic and atomic calculations of the fine and hyperfine structures of atomic hydrogen and its isotopes. As discussed by Brodsky [26, 47], this same excellent agreement between quantum electrodynamic predictions and experimental observations now occurs for all experiments which severely test the predictions.

In some respects these fine results are disappointing since there are no disagreements to point fundamental new directions for quantum electrodynamics. In other ways, however, the agreement is very gratifying. Both the theories and the experiments have been pushed so much further than once seemed possible that there may be reason to hope for still another round of increased accuracy. Furthermore, with all the troubles in theories of strong interactions, it is a pleasant contrast for quantum electrodynamics to be valid even beyond its expected limits.

5. References

[1] Lamb, W. E., and Retherford, R. E., Phys. Rev. **86**, 1014 (1952) and references therein.

[2] Applequist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970) and references contained therein.

[3] Nafe, J. E., Nelson, E. B., and Rabi, I. I., Phys. Rev. **71**, 914 (1947); **73**, 718 (1948); **75**, 1194 (1949); and **76**, 1858 (1949).

[4] Cohen, E. R., and DuMond, J. W. M., Rev. Mod. Phys. **37**, 537 (1965) and Proceedings of Conference on Physics of One and Two Electrons, 478 (1969) (North Holland Publishing Co., Amsterdam).

[5] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[6] Triebwasser, S., Dayhoff, E. S., and Lamb, W. E., Phys. Rev. **89**, 98, and 106 (1953).

[7] Robiscoe, R. T., and Cosens, B. L., Phys. Rev. **138**, A22 (1965); **168**, 4 (1968), private communications in references [2] and [5], and these Proceedings.

[8] Kaufman, S. L., Lamb, W. E., Lea, K. R., and Levinthal, M., Phys. Rev. Letters **22**, 507 (1969) these Proceedings.

[9] Shyn, T. W., Williams, W. L., Robiscoe, R. T., and Rehane, T., Phys. Rev. Letters **22**, 1273 (1969) and these Proceedings.

[10] Vorburger, T. V., and Cosens, B. L., Phys. Rev. Letters **23**, 1273 (1969) and these Proceedings.

[11] Wing, W., Ph.D. thesis, University of Michigan (1968—unpublished).

[12] Baird, J. C., Metcalf, H., Brandenburger, J., and Gondaira, K. I., Phys. Rev. Letters **21**, 165 (1968) and these Proceedings.

[13] Kleinpoppen, H., Z. Physik **164**, 174 (1961).

[14] Cosens, B. L., Phys. Rev. **173**, 49 (1968) and private communications in references [2] and [5].

[15] Lamb, W. E., and Sanders, T. M., Phys. Rev. **119**, 1901 (1960).

[16] Wilcox, L. R., and Lamb, W. E., Phys. Rev. **119**, 1915 (1960).

[17] Pipkin, F. M., Fabjan, C. W., and Brown, R. A., these Proceedings.

[18] Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. Letters **5**, 361 (1960).

[19] Bechler, R., Vessot, R. F. C., et al., IEEE (correspondence) **54**, 301 (1966).

[20] Peters, H., et al., Appl. Phys. Letters 6, No. 2, 34 (1965).

[21] Vessot, R. F. C., et al., IEEE Trans. Inst. and Meas. **IM-15**, No. 4, 165 (1966).

[22] Hellwig, H., Vessot, R. F. C., et al., private communication to be published in IEEE Trans. on Instrumentation and Measurements.

[23] Larson, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23**, 1369 (1969).

[24] Mathur, B. S., Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. **158**, 14 (1967).

[25] Parker, W. H., Taylor, B. N., and Langenberg, D. N., Phys. Rev. Letters **18**, 287 (1967); Parker, W. H., Langenberg, D. N., Denenstein, A., and Taylor, B. N., Phys. Rev. **177**, 639 (1969).

[26] Brodsky, S. J., these Proceedings.

[27] Kusch, P., Phys. Rev. **100**, 1188 (1955).

[28] Wittke, J. P., and Dicke, R. H., Phys. Rev. **103**, 620 (1956).

[29] Pipkin, F. M., and Lambert, R. H., Phys. Rev. **127**, 787 (1962).

[30] Hughes, W. M., and Robinson, H. G., Phys. Rev. Letters **23**, 1209 (1969).

[31] Crampton, S. B., Robinson, H. G., Kleppner, D., and Ramsey, N. F., Phys. Rev. **141**, 55 (1967).

[32] Low, F., and Salpeter, E., Phys. Rev. **83**, 478 (1951).

[33] Mathur, B. S., Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. **158**, 14 (1967).

[34] Avery, R., and Sachs, R. G., Phys. Rev. **74**, 433 (1948) and **74**, 1320 (1948).

[35] Sessler, A. M., and Foley, H. M., Phys. Rev. **98**, 6 (1955).

[36] Fortson, E. N., Kleppner, D., and Ramsey, N. F., Phys. Rev. Letters **13**, 22 (1964).

[37] Schwartz, C., Ann. Phys. (N.Y.) **2**, 156 (1959).

[38] Sandars, P. G. H., Proc. Phys. Soc. **92**, 857 (1967).

[39] Brenner, D., Ph.D. thesis, Harvard University (1969—unpublished).

[40] Hegstrom, R. A., Phys. Rev. **184**, 17 (1969), and these Proceedings.

[41] Grotch, H., Phys. Rev. Letters **24**, 39 (1970), private communication to be published in Phys. Rev. Letters, and these Proceedings.

[42] Hughes, W. M., and Robinson, H. G., Phys. Rev. Letters **23**, 1209 (1969) these Proceedings.

[43] Myint, T., Kleppner, D., Ramsey, N. F., and Robinson, H. G., Phys. Rev. Letters **17**, 405 (1966).

[44] Winkler, P. F., Walther, F. G., Myint, M. T., and Kleppner, D., Bulletin Am. Phys. Soc. **15**, 44 (1970).

[45] Heberle, J. W., Reich, H. A., and Kusch, P., Phys. Rev. **104**, 1585 (1956).

[46] Mittleman, M. H., Phys. Rev. **107**, 1170 (1957).

[47] Wesley, J. C., and Rich, A., Phys. Rev. Letters **24**, 1320 (1970).

[48] Zwaniger, D. E., Phys. Rev. **121**, 1128 (1961).

[49] Reich, H. A., Heberle, J. W., and Kusch, P., Phys. Rev. **104**, 1585 (1956).

[50] Ramsey, N. F., Molecular Beams, Oxford University Press (1956).

[51] Adams, E. N., Phys. Rev. **81**, 1 (1951).

[52] Greenberg, D. A., and Foley, H. M., Phys. Rev. **120**, 1684 (1960).

[53] Ramsey, N. F., Phys. Rev. **78**, 699 (1950); **87**, 1075 (1952); **90**, 232 (1953).

[54] Sternheim, M., Phys. Rev. **130**, 211 (1963).

Pressure Shifts of Hyperfine Frequencies in One-Electron Atoms*

E. S. Ensberg and C. L. Morgan

Yale University, New Haven, Connecticut 06520

Optical pumping methods for hydrogen in argon yield a pressure shift of the hydrogen hyperfine splitting $(1/\nu_0)$ $(\partial\nu_0/\partial\rho)$, which varies linearly with argon temperature (at constant density ρ) from from 5 °C to 60 °C and becomes less negative with increasing temperature by the amount of (0.20 ± 0.01) percent/degree C for hydrogen and tritium. The fractional pressure shifts for these two isotopes are $(-4.803\pm0.015)\times10^{-9}$/torr(0 °C) at about 26 °C. Work continues on deuterium. Our result for H and T suggests that the pressure shift of muonium hyperfine splitting should agree with hydrogen towith in a few percent.

Key words: Hyperfine structure; hydrogen isotopes; magnetic resonance; muonium; optical pumping; pressure shift.

1. Introduction

Collisions with inert atoms are an essential part of many high precision measurements of ground state hyperfine splittings. For example, the experiments on hydrogen, positronium and muonium described at this conference are all limited, in varying degrees, by this perturbation.

We have particular interest in this problem as it relates to the muonium measurements at Yale. Muonium is necessarily formed in a target with high gas density. The pressure (or density) shift in muonium is therefore large and, in fact, the extrapolation to zero buffer gas density dominates the quoted uncertainty in the muonium hyperfine splitting.

The muonium atom, μ^+e^-, should behave like a light hydrogen isotope in collisions with the inert gas. Therefore an experiment on the hyperfine pressure shift of hydrogen in argon should help us understand the muonium case. Our initial objectives were (a) a precise measurement of the shift in hydrogen, (b) an upper limit on the isotope dependence of the shift and (c) an upper limit for non-linear contributions to the shift, e.g., three-body collision effects. At the present time we have nearly reached these objectives but many questions about the pressure shift remain unanswered.

We will describe the experimental technique in the section that follows, and then our results for hydrogen isotopes in argon, followed by krypton and xenon. The quadratic pressure dependence is too small to measure at low pressure, so we describe another technique for measuring the shift for rubidium in argon up to 10 atm by optical pumping with white light. We will conclude with some comments on the development of a theory for the pressure shifts and implications for the muonium measurements. G. zu Putlitz[1] collaborated with us in the white light experiments. P. Crane helped develop our first hydrogen-argon apparatus.

2. Experimental Technique

Our technique is basically that of Pipkin's group in the original hydrogen optical pumping experiments [1]. Hydrogen atoms are polarized and detected by spin-exchange collisions with optically pumped rubidium vapor. Optical pumping is well adapted to pressure shift measurements because it is not excessively perturbed by the short mean free path. Our experiment differs from Pipkin's experiments in the magnetic field control, temperature control and the molecular hydrogen dissociation process. The magnetic field is critical because only field-dependent transitions are observable by the spin-exchange process. Our apparatus is set up in a remote, magnetically clean laboratory, where the earth's field is stabilized against diurnal variations by a rubidium magnetometer. Thus all of our experiments are performed at a field of about 0.569 G.

We use a sealed pyrex absorption cell containing rubidium metal, a fixed density of inert buffer gas and roughly 1 torr each of the molecular hydrogen isotopes. The cell is mounted in a plexiglass water jacket for temperature control, as shown in figure 1.

Temperature control is critical both because the frequency shift is temperature dependent and because temperature gradients produce density gradients. After months of measuring temperature gradients we concluded that only liquid thermal baths can reduce temperature differences below 1 °C across the

* Research sponsored by the Air Force Office of Scientific Research under AFOSR Grant No. 249-67.

[1] On leave from Physikalisches Institut der Universität Heidelberg, Heidelberg, Germany.

sample. We pump liquid from a constant temperature bath through the plexiglass jacket. Light pipes have recently been added to decrease the optical noise [2]. The addition of the plastic light pipes increased the temperature variation over the sample from 1/4 °C to 1/2 °C.

Dissociation of the hydrogen molecules is accomplished on the surface of a hot tunsten wire which is formed into a quadrupole array to cancel the field of the rf heating current to first order. (When the absorption cells contain tritium, we find that after initial dissociation to establish mixtures of D–T, H–T, etc., there are sufficient numbers of atoms

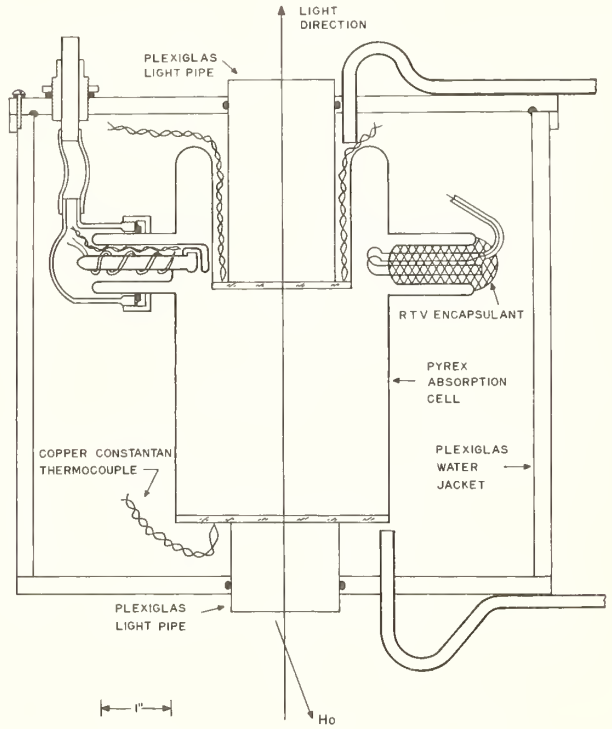


FIGURE 1. *Absorption Cell in water jacket for temperature control.*

The reentrant tubulations on each side of the upper part of the sample are shown containing a typical rubidium oven on the left and a typical quadrupole hot wire hydrogen dissociator on the right.

resulting from dissociation of the tritium to yield useable signals without further use of the hot quadrupole.)

The rubidium metal is carried in cups which can be heated by audio current in bifilial heaters to control the rubidium vapor supply over a wide range of sample temperatures. The absorption eell takes the shape of a re-entrant cylinder to exclude the hot wires and rubidium cups and their associated convection currents from the path of the transmitted light beam.

During construction of each sample, gas pressure is measured by a Texas Instruments Bourdon Tube Gauge. During the final sealoff, the sample is immersed in a room temperature water bath whose temperature is determined by NBS calibrated mercury-in-glass thermometers. The final density is known to better than 0.1 torr (0 °C). For approximate agreement with the practice in earlier experiments, we express the gas density in terms of the equivalent pressure of 0 °C: "torr (0 °C)", assuming the perfect gas law. (Unambiguous density units are badly needed in the literature of pressure shifts.)

The microwave transitions are observed by means of a silicon photo-diode in the transmitted light beam (above the sample) and are recorded by a lock-in amplifier and a chart recorder. The sum and difference of transition frequencies for the stretched states, with the Breit-Rabi formula, allows us to deduce the pressure shift of the hyperfine splitting in zero field.

3. Hydrogen in Argon

Most of our effort has been devoted to measuring the hyperfine pressure shift in argon because of its importance in muonium measurements. The resonance lines are symmetric to within 2 percent. In terms of the difference between points of maximum

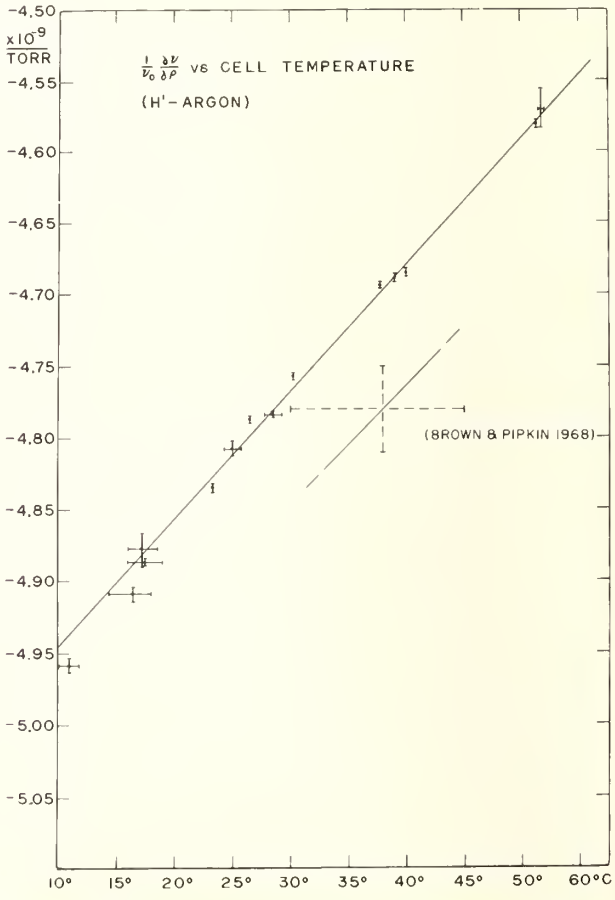


FIGURE 2. *Temperature dependence of the pressure shift for one argon sample at 124 torr (0°C).*

slope as the line width, the minimum width observed so far was 60 Hz at 220 torr. Our quoted errors are equivalent to about 2 Hz or about 1 part in 10^9 of the hyperfine frequency.

Figure 2 is a plot of the results from a single sample with an argon density of 124 torr (0 °C). Notice that the absolute magnitude of the pressure shift decreases with temperature. The temperature uncertainty is about the size of the dot except where indicated. We have similar measurements on several samples at the present time and a summary of the results appears in figure 3.

The three hydrogen isotopes have the same pressure shift at 26 °C to within our accuracy (and it is not in good agreement with muonium in argon).

We are still looking into the cause of the small difference in the deuterium and hydrogen temperature dependence. We do not believe it is a significant discrepancy. The hyperfine splitting in deuterium is only 327 MHz, so getting comparable precision is four times as difficult as in tritium and hydrogen. Our recent line widths at higher pressures correspond to an upper limit on the disorientation cross-section of less than 10^{-22} cm² for H in argon.

	$\frac{1}{\nu_0}\frac{\partial \nu}{\partial \rho}$ at 26.0°C	$\frac{\partial}{\partial T}\left(\frac{1}{\nu_0}\frac{\partial \nu}{\partial \rho}\right)$
H	$(4.803 \pm .015) \times 10^{-9}$/Torr(0°C)	$(.90 \pm .07) \times 10^{-11}$/Torr(0°C) deg C
D	$(4.803 \pm .065) \times 10^{-9}$/Torr(0°C)	$(1.23 \pm .20) \times 10^{-11}$/Torr(0°C) deg C
T	$(4.785 \pm .015) \times 10^{-9}$/Torr(0°C)	$(.92 \pm .09) \times 10^{-11}$/Torr(0°C) deg C

FIGURE 3. *Preliminary results for the hydrogen hyperfine pressure shift in argon.*

In our attempts to learn about the deuterium temperature dependence, we are in the process of applying the high temperature method that Wright, Balling and Lambert recently used in the case of deuterium in helium [3]. In their experiment, they have measured the pressure shift from near liquid air temperature to 400 °C by controlling the cell temperature with separate air blasts. So far, we have compared hydrogen and deuterium by their method up to 200 °C. We are finding that the deuterium temperature dependence is closer to hydrogen than our result in figure 3. These high temperature experiments also confirm that the shift is a monatonic function of temperature in argon, to at least 200 °C, similar to that in helium, (although of opposite sign).

4. Hydrogen in Krypton and Xenon

Our first published results from this experiment were for hydrogen in krypton and xenon. We wanted to measure the pressure shift in krypton and xenon [4] before the first muonium measurements in krypton. Our method was similar, without a water bath. The results were limited by signal-to-noise problems because of large depolarization cross sections in these gases. We found that the fractional pressure shift $\nu_0^{-1}[(\partial/\partial\rho)\nu] = (-10.4 \pm 0.2) \times 10^{-9}$/torr (0 °C) in krypton and $(-20 \pm 2) \times 10^{-9}$/torr (0 °C) in xenon. No temperature effect was observed because the measurements were made in a narrow temperature region. The shift for muonium in krypton turned out to be in good agreement with our result.

5. High Pressure Optical Pumping by White Light

The results thus far have been linear in pressure. Since low pressure experiments are very insensitive to a quadratic dependence on pressure, we extended

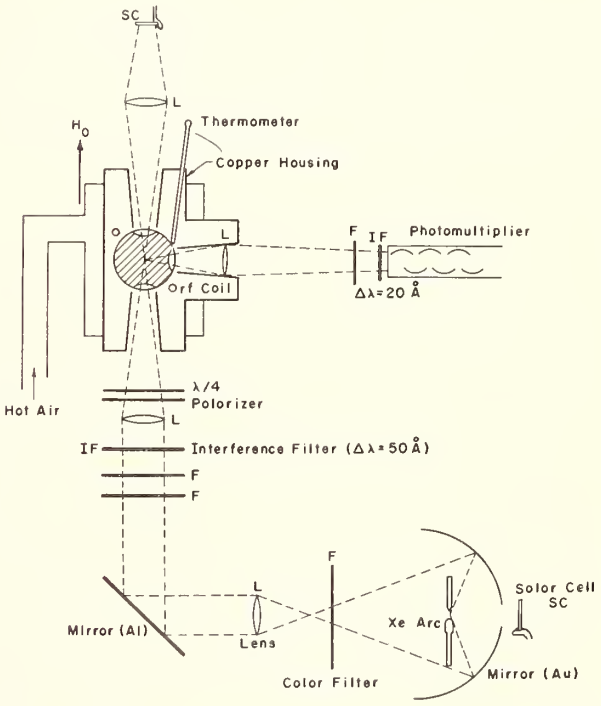


FIGURE 4. *Apparatus for optical pumping with white light.*

the technique to measurements on rubidium in argon at 10 atm [5]. At high pressure, the rubidium optical line is broadened by about 0.4 Å per atm [6], so that even the low spectral density of an arc lamp becomes more efficient than a resonance lamp when the pressure in the absorption cell exceeds about 1 atm. So we replaced our rubidium lamp with a 1 KW Xe arc lamp (see fig. 4). Then we reduced the white spectrum to about 50 Å by a series of filters centered on 7948 Å to avoid heating and instrumental scatter. All our signals were detected in fluorescence to avoid the 90 percent of the light not involved in the pumping process. Our absorption cells were well annealed pyrex and were each tested hydraulically at 50 atm before use. Expressing the pressure shift

as a fraction $(\nu-\nu_0)/\nu_0=a\rho+b\rho^2$, we found $a=(-11.5\pm0.3)\times10^{-9}$/torr (0 °C), $b=(+10\pm4)\times10^{-14}$/torr (0 °C) for argon at 80 °C. The shift grows more negative with increasing temperature at (0.60±0.04) percent/°C and this temperature dependence seems linear within these limits between 70 °C and 130 °C. There is also no disagreement between ^{85}Rb and ^{87}Rb to within 0.1 percent. This result is in agreement with measurements by Bender, Beaty and Chi [7] assuming a linear extrapolation in temperature to 21 °C. The quadratic result is roughly compatible with the upper limit detectable in the muonium experiments. This quadratic shift is about the sign and magnitude one would expeet from the simple hypothesis that the major effect of the three-body collisions is to shield the seeond argon atom from a collision with the hydrogen atom.

The white light technique may have other applications for high precision measurements. For example, in Rb frequency standards, light shifts of the dispersion type [8] (which depend on the displacement of structure in the lamp spectrum from the center of the absorber line) might be reduced by the broad spectrum of a white-light source.

6. Theory

The theory so far published on the hyperfine pressure shift all starts with the idea of Adrian [9] for H atoms frozen in inert gas lattices. He assumed the shift is the result of a competition between two static interactions: a long range attraction force between the H atom and its nearest neighbors which extends the hydrogen wave function, reducing the hyperfine contact interaction at the nucleus, and a short range repulsive interaction which compresses the wave function and increases the hyperfine interaction. Adrian thus accounted for both negative pressure shifts in heavy inert gases and positive pressure shifts in light inert gases.

Margenau and Herman [10] applied the statistical theory of spectral lines to this static interaction to derive the frequency shift for the gaseous state. Usually phenomenological potentials must be used but for helium, however, precise calculations of this model have been made by Clarke [11] and T. P. Das and eo-workers [12]. Improvements [13] in the variational calculation for H and He have gradually brought the theory and the experiment together on both the magnitude of the shift at 50 °C and apparently also on the temperature dependence [14]. The significance of this agreement is not entirely clear since the model is essentially static. The good agreement does not follow in other cases: Weiss, Balling, and Lambert found the temperature dependence of nitrogen in He in disagreement with the theory [15, 16].

No dependence on the hydrogen isotope is predicted by these theories if the classical Boltzmann distribution is used in the statistical theory. However, Clarke replaced the Boltzmann distribution by a quantum statistical distribution, which adds a term inversely proportional to the redueed mass of the He–H system. The size of this term has not been estimated but it suggests that a linear extrapolation from our experimental upper limit on the isotope effect for H and T to the muonium atom might be valid.

7. Conclusions

Finally, then, we have measured the pressure shift as a function of temperature of H in argon and we have compared it with tritium (and, with less success right now, deuterium).

The agreement between T and H is sufficiently close to conclude that if the present theory of the collisions were correct, muonium should agree with H to within about 3 pereent. Actually, the results of the Yale group [17] indicate that the fractional hyperfine pressure shift is low by about 20 percent.

Furthermore we have measurements for H in Kr and xenon which indicate the pressure shift increases far more rapidly than the polarizability for those atoms. This failure is not surprising in view of the importance of molecular formations in the collisions, as recently demonstrated by the relaxation measurements in heavy inert gases by Bouchiat [18]. It is clear that to understand these collisions phenomenologically, at least, it will be necessary for us to perform similar experiments on hydrogen. We have not yet done this.

On the basis of our results for the temperature dependence, we can speculate that if the muonium atoms should happen to have an average temperature of about 100 °C, they would agree with hydrogen. Actually, it is difficult to think of simple reasons why muonium atoms should not have the same temperature as the target gas (about 25 °C) after only a nanosecond or so. This unlikely speculation is made even more so by the observation of a *larger* fractional pressure shift at lower pressure by the Chicago group [19].

8. Acknowledgments

We wish to acknowledge the support and encouragement of Professor V. W. Hughes in this work, and to thank Professor R. L. Wolfgang for the use of his tritium-handling facilities. Mr. E. C. Brosious worked out the glass-blowing problems and Mr. R. J. Broughton built essential parts of the electronics.

9. References

[1] Anderson, L. W., Pipkin, F. M., and Baird, J. C., Jr., Phys. Rev. **120**, 1279 (1960); Phys. Rev. **121**, 1864 (1961); Phys. Rev. **122**, 1279 (1960); Brown, R. A., and Pipkin, F. M., Phys. Rev. **174**, 48 (1968).
[2] Bedard, F. D., Private communication.
[3] Wright, J. T., Balling, L. C., and Lambert, R. H., Phys. Rev. **1A**, 1018 (1970).
[4] Ensberg, E. S., and Morgan, C. L., Phys. Letters **28A**, 106 (1968).

[5] Ensberg, E. S., and zu Putlitz, G., Phys. Rev. Letters **22,** 1349 (1969).
[6] Chén, S. Y., Phys. Rev. **58,** 1051 (1940).
[7] Bender, P., Beaty, E. C., and Chi, A. R., Phys. Rev. Letters **1,** 311 (1958).
[8] Barrat, J., and Cohen-Tannoudji, C., J. Phys. et Rad. **22,** 329 (1961).
[9] Adrian, F. J., J. Chem. Phys. **32,** 972 (1960).
[10] Herman, R., and Margenau, H., Phys. Rev. **122,** 1204 (1961).
[11] Clarke, G. A., J. Chem. Phys. **36,** 2211 (1962).
[12] Ray, S., Lyons, J. D., and Das, T. P., Phys. Rev. **174,** 104 (1968); Phys. Rev. **181,** 465 (1969).
[13] Das, G., and Ray, S., Phys. Rev. Letters **24,** 1391 (1970).
[14] Rao, B. K., and Das, T. P., Phys. Rev. **185,** 95 (1969).
[15] Weiss, R. E., Balling, L. C., and Lambert, R. H., Bull. Am. Phys. Soc. **15,** 489 (1970).
[16] Ray, S., and Das, T. P., Phys. Rev. **174,** 32 (1968); Phys. Rev. **181,** 465 (1969); Lyons, J. D., Ray, S., and Das, T. P., Phys. Rev. **174,** 112 (1968); Phys. Rev. **181,** 465 (1969).
[17] Crane, P., Amato, J., Hughes, V. W., Lazarus, D. M., zu Putlitz, G., and Thompson, P. A., Bull. Am. Phys. Soc. **15,** 45 (1970).
[18] Bouchiat, C. C., Bouchiat, M. A., and Pottier, L. C. L., Phys. Rev. **181,** 144 (1969).
[19] Ehrlich, R. D., Hoffer, H., Magnon, A., Stowell, D., Swanson, R. A., and Telegdi, V. L., Phys. Rev. Letters **23,** 513 (1969).

Measurement of the Muon Magnetic Moment by "Double Resonance" in Muonium, and a New Value of Alpha*

A. Magnon,** R. DeVoe,*** P. M. McIntyre, D. Y. Stowell, R. A. Swanson,**** and V. L. Telegdi

The Enrico Fermi Institute and Department of Physics, The University of Chicago, Chicago, Illinois 60637

Heretofore, any determination of α from the hfs splitting $\Delta\nu$ of muonium was severely limited by uncertainties in the published magnetic moment (not g-factor!) $\mu_\mu \equiv \mu_B g_\mu'$s of the muon. We have now determined μ_μ from the frequencies of the two alternative Zeeman transitions ν_1 (F=1, M=1↔F=1, M=0) and ν_2 (F=1, M=−1↔F=0, M=0) at that "magic" field B_0 (∼11.35 kG) value where $\partial\nu_1/\partial B=\partial\nu_2/\partial B=0$. At B_0 one has $(\nu_2-\nu_1)/(\nu_2+\nu_1)\equiv\nu_-/\nu_+=2[G(1-G)]^{1/2}$, $G=-g_\mu'/(g_j-g_\mu')$. Direct resonances in ν_- (or ν_+) were obtained by exciting the cavity simultaneously with ν_1 and ν_2 and keeping ν_+ (or ν_-) fixed. We find $g_j/g_\mu'=206.76509$ (80), which with $g_j=g_e(1-\alpha^2/3)$ corresponds to $f_\mu/f_p=3.183373$ (13); pressure shifts of g_j are expected to be negligible here. This result, combined with our best value for $\Delta\nu$ (extrapolated to zero pressure), yields $\alpha^{-1}-137=0.03654$ (30) (2.0 ppm), in good agreement with the TPL recommended value 0.03602 (21).

Key words: Fine structure; muonium hfs; muon magnetic moment.

1. Introduction

Knowing both the ground-state hyperfine interval $\Delta\nu(\mu e)$ and the muon magnetic μ_μ to high accuracy (say, <5 ppm) one can—aside from checking the QED corrections to $\Delta\nu(\mu e)$—do two interesting things; (a) determine, with competitive precision, the fine-structure constant α, from an independent new source, (b) set limits on the polarizability of the proton from the observed ratio $\Delta\nu(pe)/\Delta\nu(\mu e)$. To date this program could unfortunately not be realized in practice. While $\Delta\nu(\mu e)$ has become known with increasing accuracy over the past decade, (from 13 ppm in 1964 [1] to 2.0 ppm now) there has been no corresponding progress in the measurement of μ_μ. The most recent published determination [2] of μ_μ, has an accuracy of 9 ppm, based on a comparison of the muon and proton precession frequencies in water [3]. This determination, as well as any other based on muon precession in matter, furthermore requires a systematic correction, presumably of the same magnitude [4], to allow for differences in diamagnetic shielding.

We describe here a determination of μ_μ and $\Delta\nu$ from the frequencies of the Zeeman transitions $(F, M_F)(1, 1)\leftrightarrow(1, 0)$ and $(1, -1)\leftrightarrow(0, 0)$ in the region of intermediate coupling [5]. The Breit-Rabi formula yields for these

$$\nu_{1(2)}=(\Delta\nu/2)[1-(+)D(x, G)], \qquad (1)$$

where

$$D=(1+x^2)^{1/2}-x(1-2G),$$

with

$$G=g_\mu'(M)/[g_\mu'(M)-g_j]=g_\mu'/(g_\mu'-g_e),$$

$$g_\mu'(M)=(1-\alpha^2/3)g_\mu',$$

$$g_\mu'=(m_e/m_\mu)g_\mu, \quad \text{and} \quad x=B\mu_0(g_\mu'(M)-g_j)/h\Delta\nu.$$

We choose to work at that "magic" field $x_0(G)$ where $\partial D/\partial x=0$, i.e., where both ν_1 and ν_2 become, to first order, field-independent. The advantages of this choice have already been discussed [6] in connection with a measurement of ν_1. At x_0 ν_1 and ν_2 depend obviously, aside from proportionality of $\Delta\nu$, on G along, as do their sum and difference. Explicitly, one has

$$\nu_+=\nu_1+\nu_2=\Delta\nu \qquad (2a)$$

$$\nu_-(x_0)=\nu_2-\nu_1=2\Delta\nu[G(1-G)]^{1/2} \qquad (2b)$$

$$\nu_-/\nu_+=2[G(1-G)]^{1/2}. \qquad (2c)$$

Note that the ratio of direct interest (2c) for determining g_μ' is not affected by the pressure dependence of $\Delta\nu$.

While it would be adequate to determine ν_1 and ν_2 in independent runs (at x_0), there are statistical and systematic advantages to measuring ν_+ and ν_- *directly by "double resonances,"* as we shall now discuss. Inasmuch as stopping (polarized) muons can capture electrons of *either* spin direction, muonium is formed 50–50 to one of the upper and lower

* Research supported in part by NSF Grant No. GP 6035.
** N. S. F. trainee, 1969/70.
*** Visitor from Centre à l'Energie Atomique, Saclay, France.
**** On sabbatical leave from the University of Calif. at San Diego, LaJolla.

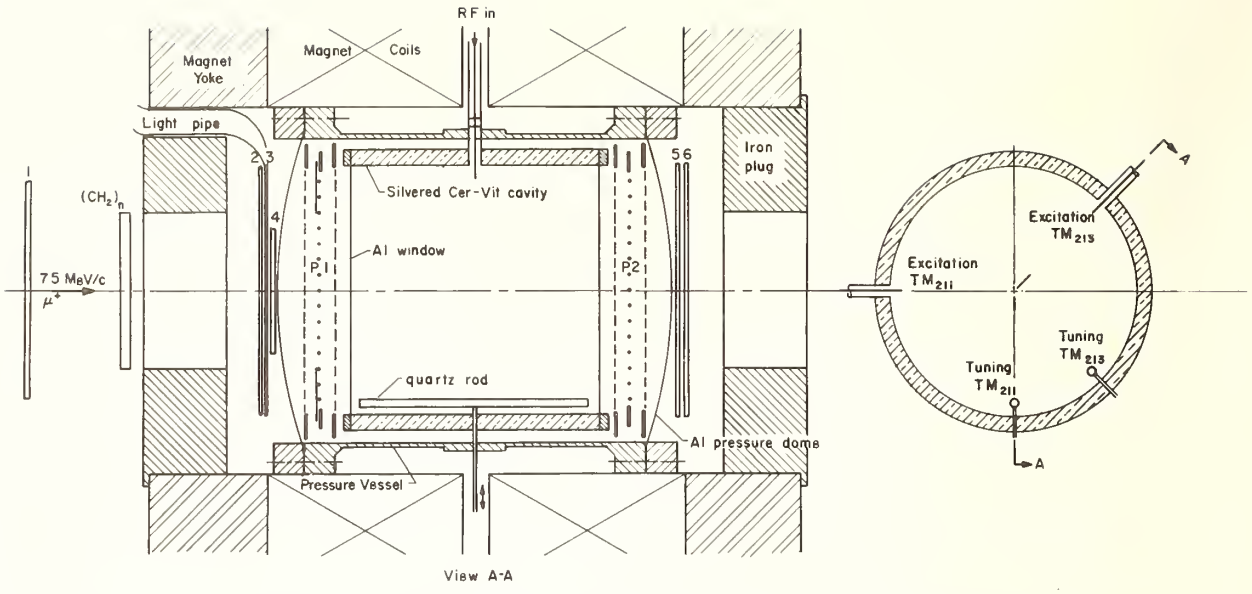


FIGURE 1. *Experimental set-up to observe "double resonances" in muonium.*
The two iron plugs provide a better homogeneity of the axial static field. The RF modes have been rotated at 45° to ensure decoupling.

Zeeman levels, for example (F = 1, M = 1; F = 0, M = 0). Thus inducing ν_1 or ν_2 transitions is completely equivalent. By having *two* microwave fields of the appropriate frequencies $\nu_{1,2}$ simultaneously present, one can hence induce a resonant spin flip transition for each (μe) atom formed. By keeping $\nu_+(\nu_-)$ constant, one can thus sweep out $\nu_-(\nu_+)$; the resultant "double resonance" curves will have twice the height and width that a single resonance in ν_1 (or ν_2), would exhibit, provided one chooses the two rf power levels so as to insure equal signals in the two single (ν_1, ν_2) resonances. The importance of this proviso and of the initial choice of the frequency that is kept constant will be discussed later.

2. Experimental Arrangement

The apparatus used was substantially that of reference [6], except for the following essential modifications: (a) the field homogeneity was greatly improved (to $<0.1\%$ over the relevant volume) by reducing the end apertures of the magnet with iron plugs (see fig. 1); (b) the microwave cavity (made of CerVit) was dimensioned for simultaneous excitation of the TM_{211} ($\nu_1 \simeq 1920$ MHz) and TM_{213} ($\nu_2 \simeq 2540$ MHz) modes. An azimuthal maximum of one of these corresponds to a minimum of the other, so that both modes can be tuned independently with two suitably placed quartz rods. Microwave power was provided by two klystrons, each driven by a thermostated, crystal-controlled oscillator. During each run, the rf powers were kept stable to ± 1 percent and the frequencies to ± 0.01 ppm; the choice of decoupled modes enabled one to measure their Q's independently.

3. Experimental Procedure and Results

First, single resonances in ν_1 and ν_2 as in reference [6] were taken, with rf power levels so chosen as to produce essentially identical signals; minor power adjustments were then made to make the observed curves equal. Next, double resonances were performed under the same conditions, varying ν_1 and ν_2 so as to keep either their sum or difference constant (at "best estimate" values based on existing knowledge of $\Delta\nu$, g_μ' and $\partial\Delta\nu/\partial p$). Figure 2 shows a typical ν_+ resonance so obtained. Inasmuch as the field-dependence term $D(x)$ affects ν_1 and ν_2 with opposite signs, it cancels in their sum, and one performs in

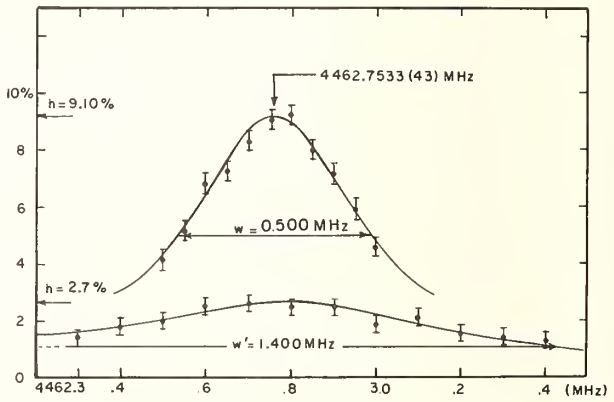


FIGURE 2. *The upper curve is one of the double resonances for the sum of frequencies at 11.830×10^3 torr of krypton.*
The transitions are (F, M_F); $(1, 1) \leftrightarrow (1, 0)$ and $(1, -1) \leftrightarrow (0, 0)$. Error bars are counting statistics (1 standard deviation). The lower curve (ref [6]) is a resonance for the same frequency obtained at 0.010 G in 32.800×10^3 torr of krypton. It should be noticed that for a given counting rate, the error on the fitted center is for a Lorentzian line proportional to W/h.

measuring ν_+ essentially a direct field-independent determination of $\Delta\nu$, as in zero field. We therefore reproduce in figure 2 for comparison an actual "zero field" resonance [7].

The quadratic and higher field corrections add however for ν_-, and it is hence for this resonance particularly important to work at the proper field. Fortunately, the double resonance in ν_- *can itself be exploited to establish the correct "magic" field.* One simply measures $\nu_-(x)$ in the neighborhood of the "best estimate" for x_0, and fits a parabola (actually a cubic) of *a-priori* known curvature to the experimental points. The parameters of this fit are essentially decoupled, so that one obtains $\nu_-(x_0)$ without significant loss in statistical power. Note that this procedure, illustrated in figure 3, automatically yields the correct *effective* average field, properly

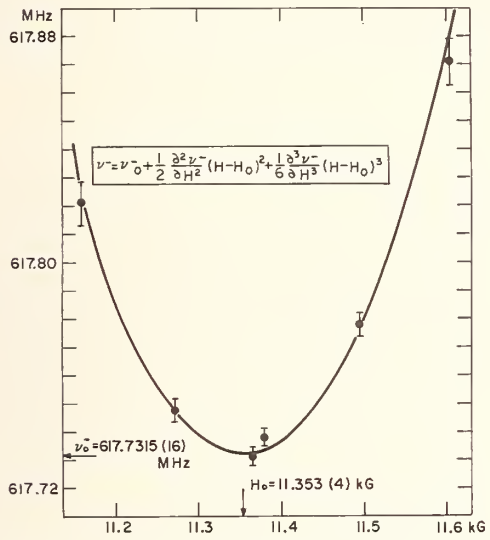


FIGURE 3. *Difference frequency ν_- versus static field.* Errors (1 standard deviation) include counting statistics and field uncertainties.

weighted with the muon stop distribution, positron detection efficiency, etc.

The main series of measurements was done in krypton at fairly high pressure (230 psi at 30 °C), interspersing ν_- and ν_+ runs at the same gas density. Additional measurements of ν_+ were performed in Kr and Ar at low pressure for extrapolation purposes. The results are as follows:

$$\nu_+ \text{ (Kr, 11 830 torr)} = 4462.7504\ (22)\ \text{MHz} \quad (3a)$$

$$\nu_- \text{ (Kr, 11 830 torr)} = 617.7299\ (12)\ \text{MHz} \quad (3b)$$

$$\nu_+ \text{ (Kr, 2598 torr)} = 4463.1824\ (50)\ \text{MHz} \quad (3c)$$

$$\nu_+ \text{ (Ar, 3030 torr)} = 4463.220\ (22)\ \text{MHz}, \quad (3d)$$

where the numbers in brackets indicate the "equivalent densities" [8] in torrs at 0 °C, while the errors are standard deviations allowing for systematic uncertainties. Using eq (2c) and assuming $g_j = g_e(1-\alpha^2/3)$, one obtains from (3a, b)

$$g_e/g_\mu' = 206.76509\ (90)\ (4\ \text{ppm}) \quad (4)$$

which is equivalent to

$$f_\mu/f_p = 3.183373\ (13)\ (3.5\ \text{ppm}). \quad (4')$$

This latter ratio may be compared to $f_\mu/f_p = 3.18336\ (3)$, based on a direct measurement [2] of f_μ; the agreement with (4'), while not statistically compelling, suggest that no correction of the magnitude advocated by Ruderman [4] is necessary.

Extrapolating (3a) with (3c) to zero density one obtains

$$\text{Kr, } \Delta\nu(0) = 4463.3040\ (100)\ \text{MHz} \quad (5a)$$

corresponding to a fractional pressure shift (FPS): $(1/\Delta\nu)(\partial\Delta\nu/\partial p) = 10.50\ (0.32) \times 10^{-9}$ torr^{-1}. This FPS agrees with that found for H in Kr $(10.4\ (0.20) \times 10^{-9}$ torr^{-1} [9]); the extrapolation of (3c) with the latter shift yields

$$\text{Kr, } \Delta\nu(0) = 4463.3030\ (62)\ \text{MHz} \quad (5b)$$

which is of course consistent with (5a). Similarly, (3d) extrapolates using a FPS $= 5.44\ (45) \times 10^{-9}$ torr^{-1} [6] to

$$\text{Ar, } \Delta\nu(0) = 4463.293\ (23)\ \text{MHz} \quad (5c)$$

which incidentally confirms an earlier result [5]. Thus the Kr and Ar values for $\Delta\nu(0)$ appear consistent with each other, and we take as the final value

$$\Delta\nu(0) = 4463.3022\ (89)\ \text{MHz} \quad (2.0\ \text{ppm}). \quad (5d)$$

Using the theoretical expression [1] for $\Delta\nu$, one obtains with (4) and (5d)

$$1/\alpha = 137.03654\ (30) \quad (2.0\ \text{ppm}) \quad (6)$$

which agrees well with the currently "recommended" [10] value, $1/\alpha = 137.03602\ (21)$. Conversely, this agreement can be interpreted as a quantitative experimental verification of the QED corrections [11] entering $\Delta\nu$.

The ratio of the observed hydrogen and muonium hf intervals was formerly used [6] to obtain (μ_μ/μ_p) assuming that the proton polarizability ϵ is negligible. Now that this moment ratio is measured (4), one can, using (5d) and $\Delta\nu(pe) = 1420.40575$ [12] (5), derive the value

$$\epsilon = (-9.6 \pm 4)\ \text{ppm}. \quad (7)$$

4. Systematic Errors

There are two working hypotheses underlying our method for measuring ν_- (or ν_+); (a) that the frequency not being determined, i.e., $\nu_+(\nu_-)$ is main-

tained at its *correct* (and unknown!) value ν_+^0 (or ν_-^0); (b) that the signals contributed are *identical*. Obviously both hypotheses are only approximations to experimental reality, and we have to discuss the systematic errors induced by them. The observed signal will have the general form

$$L(\nu_+) = L_1[(\nu_+ - \nu_+^0 + \delta\nu_-)^2] + L_2[(\nu_+ - \nu_+^0 - \delta\nu_-)^2], \quad \text{where} \quad \delta\nu_- \equiv \nu_-^0 - \nu_-$$

(i.e., the departure of ν_- from its correct value). Thus the center of a symmetric (say Lorentzian) curve fitted to the data will in general be shifted with respect to ν_+^0, precisely by a systematic uncertainty Δ_+. This uncertainty, which is proportional to $\Delta\nu_-$, depends on the differences between L_1 and L_2 (as measured, see above); a numerical study gives

$$\Delta_+ \simeq \delta\nu_-/6.$$

Thus, e.g., in the measurement yielding (3a) the uncertainty $\delta\nu_- = 6$ kHz in the set value of ν_- ($= 617.732$ MHz) contributed $\Delta_+ = 1$ kHz to the quoted error. The situation for ν_- is entirely analogous, except that to obtain (3b) ν_+ was fixed at its best estimate at this time yielding $\Delta_- = 0.85$ kHz.

Another source of potential systematic error is the so-called g_j-shift, i.e., the fact that not only $\Delta\nu$ but also g_j can be affected by collisions with the host gas atoms [12]. For Rb in He, where this phenomenon has been measured [13], the effect is small. It is not clear how one should extrapolate to H in Kr, a case for which no direct measurement exists. Considering the theory of effect, in particular eq (15) of [14], one would conclude that the g_j-shift in our case should be negligible.

5. Acknowledgments

As in our previous work on muonium [5], this experiment was made possible through the continued generosity of many individuals and companies. We are particularly indebted to Dr. E. L. Ginzton (Varian Associates) for a 4K3SK Klystron and two precision oscillators, and to Mr. J. Duncan (Owens-Illinois) without whose enthusiastic cooperation our CER-VIT cavity could not have been obtained. We are grateful to M. Neumann, R. Norton, and T. A. Nunamaker of this Institute for their unfailing technical assistance, and wish to thank Dr. H. G. Robinson (Duke), Dr. E. Ensberg (Yale) and Dr. R. M. Herman (Pennsylvania State) for several illuminating discussions.

Note added in proof: Since this paper was presented, several new relevant facts have become known, viz:

(a) A new measurement of f_μ/f_p has been published [J. F. Hague et al., Phys. Rev. Letters **25**, 628 (1970), and these Proceedings]. Its result, $f_\mu/f_p - 3.183\,347$ (9), agrees with (4′) to 2σ and yields with our $\Delta\nu$ essentially the "recommended" α.

(b) Fulton, Owen, and Repko (Phys. Rev. Letters **26**, 61 (1971)) have completed a new QED correction term to $\Delta\nu(\mu e)$ which raises this frequency by 5.4 ppm. This entails an increase in α^{-1} by 2.7 ppm for given values of μ_μ and $\Delta\nu$!

(c) Using semi-empirical extrapolation, Dr. R. M. Herman has now suggested that g_j(M) should suffer a pressure shift of -11 ppm under our experimental conditions. Assuring this g_j shift and using the corrected $\Delta\nu$ formula of (b), we find $f_\mu/f_p = 3.183\,337$ (13) in perfect agreement with (a), and $\alpha^{-1} = 137 = 0.03617$ (30), where the errors contain no allowance for uncertainties in the g_j shift.

6. References

[1] Hughes, V. W., Annual Reviews of Nuclear Science **16**, 445 (1966).
[2] Hutchinson, D. P., Larsen, F. L., Schoen, N. C., Sober, D. I., and Knofsky, A. S., Phys. Rev. Letters **24**, 1254 (1970).
[3] More accurate precision measurements of $f\mu$ are discussed in Bull. Am. Phys. Soc. II **15**, 608 (1970), paper HK7, but detailed results are as yet unpublished.
[4] Ruderman, M., Phys. Rev. Letters **17**, 794 (1966).
[5] There is considerable similarity between this work and the g_e/g_p determination of Myint, T., et al., Phys. Rev. Letters **17**, 405 (1966). Our notation (M) for *bound-state* g-factors is analogous to their (H).
[6] Ehrlich, R. D., Hofer, H., Magnon, A., Stowell, D., Swanson, R. A., and Telegdi, V. L., Phys. Rev. Letters **23**, 513 (1969).
[7] Thompson, P. A., Amato, T. T., Crane, P., Hughes, V. W., Mobley, R. M., zu Putlitz, G., and Rothberg, J. E., Phys. Rev. Letters **22**, 163 (1969).
[8] Kr densities are expressed in terms of the pressure at 0 °C, using the virial data of Beattie, J. S., Brierley, J. S., and Bariault, R. J., J. Chem. Phys. **20**, 1615 (1952).
[9] Ensberg, E. S., and Morgan, C. L., Phys. Letters **28A**, 106 (1968).
[10] Parker, W. H., Taylor, B. N., and Langenberg, D. N., Rev. Modern Phys. **41**, 375 (1959).
[11] Drell, S. D., and Sullivan, J. D., Phys. Rev. **154**, 1477 (1967).
[12] Crampton, S. B., Kleppner, D., and Ramsey, N., Phys. Rev. Letters **11**, 338 (1963).
[13] This possible effect was first mentioned to us by Dr. C. Bouchiat.
[14] Hayne, G. S., and Robinson, H. G., private communication.
[15] Herman, R. M., Phys. Rev. **136A**, 1576 (1964).

DISCUSSION

A. Schenck: I would like to ask one question. Did you assume the same pressure shift for both high-frequency transitions for the hyperfine structure. Is the pressure shift the same for both transitions or not?

A. Magnon: No, the pressure shift for muonium in argon is not the same as the pressure shift for krypton.

W. E. Lamb, Jr.: You have transitions ν_1 and ν_2.

The question is whether the pressure shifts are the same for those two transitions.

A. MAGNON: The measured quantities are not ν_1 and ν_2. They are the sum $\nu_1+\nu_2$ and the difference $\nu_2-\nu_1$. Yes the pressure shift is the same for every value—

A. SCHENCK: Do you have some proof of this?

A. MAGNON: No, I don't see how it could be—I don't quite see your point.

W. E. LAMB, JR.: My guess would be that the zero field splitting ΔE would be pressure dependent, and for the rest you could rely on the Breit-Rabi formula. But, of course if the g values were affected by collisions, that would make a difference.

A. MAGNON: Yes, that's the only possible pressure shift and something that has not been measured, but we have done an extrapolation using measurements of the g_j shift, for rubidium in argon, and in krypton, and we found that it's extremely small. It's much less than a tenth of a ppm.

Precision Measurement of the Magnetic Moment of the Muon*

J. F. Hague, J. E. Rothberg, A. Schenck, D. L. Williams,
R. W. Williams, and K. K. Young

Physics Department, University of Washington, Seattle, Wash. 98105

and

K. M. Crowe

Lawrence Radiation Laboratory, University of California, Berkeley, Calif. 94720

The ratio of muon to proton magnetic moment (μ_μ/μ_p) has been measured to high precision in three chemical environments. The agreement shows that the "Ruderman correction" is not applicable. The result is $\mu_\mu/\mu_p = 3.183347(9)$ (2.8 ppm); in terms of the muon mass, this implies $m_\mu/m_e = 206.7683(6)$.

Key words: Magnetic moment of muon; muon.

The ratio of muon to proton magnetic moment, $\mu_\mu/\mu_p = (g/m)_\mu/(g/m)_p$ is needed to extract the muon's anomalous magnetic moment, $(g-2)/2$, from the observed frequency in a "$g-2$" experiment, $\omega_{g-2} = (g-2)eB/2$ mc. Of more immediate interest, it enters in the relation between the muonium hyperfine splitting, ν_m, and the fine structure constant, α. The three most recent measurements, which have errors of 13 to 22 parts per million (ppm), are not sufficiently precise to take advantage of the accurate muonium results now available [4, 5]. Ruderman [6] suggested that the substantial discrepancy between α determined from hydrogen hfs and from the then-current muonium hfs and the Columbia value [1] of μ_μ/μ_p could be partially reconciled by applying to the latter a chemical correction amounting 15 ppm. We report [7] new high-precision measurements of μ_μ/μ_p which are 10 ppm below the Columbia result; we show that the Ruderman correction [6] is not applicable; and we find that two newly reported muonium results [4, 5] bracket the value for ν_m predicted by our ratio and the currently accepted value of α.

The method is to use the muon decay asymmetry to observe the precession frequency, geB/2 mc, of a sample of polarized positive muons at rest in a magnetic field, and to observe the resonance frequency of protons in the same field. A 200-MeV/c muon beam was obtained from pions produced at the LRL 184-inch cyclotron. Figure 1 represents the arrangement of counters and target in the magnet. The stopped-muon logic was

$$(\text{Beam})\,\text{HM}_{\text{dynode}}\text{S1}\overline{\text{S2A1A2}},$$

and the decay electron was

$$\text{Se}\,(\text{E1 or E2})_{\text{dynode}}\text{S4}\overline{\text{S1A1A2M}}_{\text{dynode}}.$$

Timing signals from the muon counter M and the electron counters E were presented to fast discriminators with thresholds set 1/4th the trigger thresholds; the output signals were then passed by gated discriminators that were gated on (in a few nanoseconds) if the logic requirements had been met. These gated timing signals then opened (M) and closed (E) the gates of fast scalers which scaled a free-running oscillator. The timing between the muon and electron signals was done by two independent systems: a "digitron" with an effective least count of 1.25 ns obtained from a 400-MHz clock and two suitably phased scaling systems, and a Hewlett-Packard timing counter (HP5360A) based on a 10-MHz clock and internally converted analogue interpolation. The digital information on each event included the two time interval measurements and records of extra counts which could affect the data: second counts in either the E channel or the M channel during the time the gate was open, and any count in an E counter during the 5 μs preceding the gate opening. This information was stored by an on-line computer, and every few seconds was transferred to magnetic tape along with the digital record of the proton NMR frequency of the "monitor" probe. Details of the method and the many checks

* Work supported by the NSF and the U.S. Atomic Energy Commission.

on the system will be published elsewhere. The most important point is that the elapsed time for each muon-electron event is recorded with a simple and direct method by counting cycles of a free-running crystal-controlled oscillator. Such a system was used on the muon $g-2$ experiment; [8] it has many internal checks, and can be made highly reliable.

The accumulated data represent the number of events versus elapsed time; it is an exponential, modulated with the frequency we seek. Figure 2 shows a part of the data for one stopping substance.

The stopping material was liquid in a 3-in cube made of 5-mil Mylar. The container accounted for 1 percent of the total counting rate; the target-out rate was 2.5 percent. The decay-electron rate was 60/s, with an asymmetry of 0.16 in water. The target-out asymmetry was 0.05. We measured the background frequency and corrected for it. A large bending magnet with special pole tips and shimming coils [9] gave a field with weighted average of 0.3 ppm above the value at the center of the gap, and rms deviation of 2 ppm. The field was 11 kG, corresponding to about 149 MHz for muons and 46.8 MHz for protons.

Two separate proton magnetic resonance systems were used; one was part of the magnetic-field regulation and the other served to monitor the field during running (as shown in fig. 1) and to map the field (four field maps were made). Proton resonance was observed in a small cylindrical sample of H_2O+ 0.005 M $Fe(NO_3)_3$. Frequency at the monitor position was continuously recorded by a crystal-controlled counter. The field at the center of the gap (target out) was measured every few hours. A small bulk-susceptibility correction was made because the NMR sample and the stopping volume do not have the same shape.

The correct average over the magnetic-field map involved an auxiliary experiment. The stopping distribution and decay asymmetry were measured as functions of position in the stopping volume, and the final weight at each point was the product of asymmetry and counting rate.

We have made measurements in NaOH solution, distilled water, and methylene cyanide, $CH_2(CN)_2$. A maximum-likelihood fit was made to the data from each of the two timing systems, leaving frequency, phase asymmetry, and (uniform) background as free parameters [10]. The frequency was determined, in each case, to about 2 ppm statistical accuracy. The overall agreement between results from the two independent systems was 0.5 ppm. Starting or ending the analysis interval at different times had no significant effect. The corrections and systematic errors are summarized in table 1.

Results are in table 2. We see no significant difference between NaOH solution and distilled water. The effect suggested by Ruderman [6] requires the presence of the muon as a positive ion. However, OH^- is known to recombine with H^+ in water at an extremely rapid rate, and it can be shown [11] that the μ^+ ions would become neutralized, in 0.1 N NaOH solution, in $<10^{-10}$ s. The frequency in NaOH solution, expected according to Ruderman to be $\simeq$15 ppm lower than in H_2O, is in fact 1.6 ppm higher.

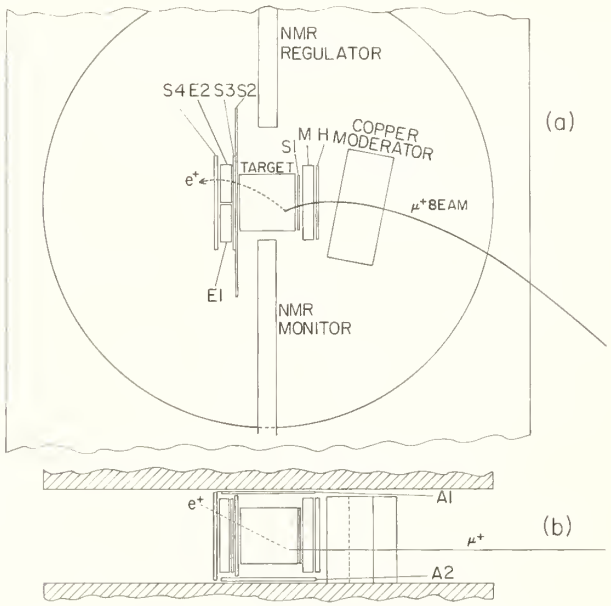


FIGURE 1. *Plan view (a) and elevation (b) of the apparatus.* The 24-in circle is a special pole-tip assembly fitting inside the 29×36 in main gap of the magnet. Collimators, etc., have been omitted.

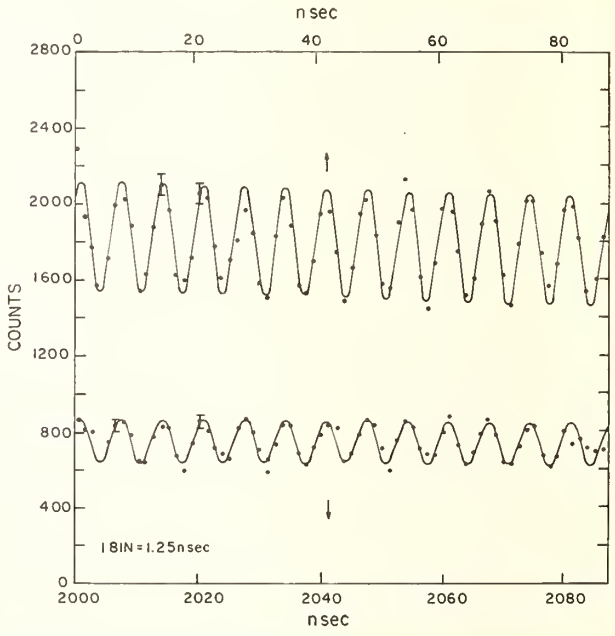


FIGURE 2. *Two short sections of the data for one target material* (0.5 N NaOH), *for which there were 3.4 million analyzed events.*
The smooth curve is the maximum-likelihood fit.

TABLE 1. *Corrections to ω_μ/ω_p, and systematic-error assignements*

Effect	Correction (ppm)	Error (ppm)
Proton resonance frequency at magnet center		0.9(H_2O)
		1.3($CH_2(CN)_2$)
Weighted average over field map		1.0
Bulk susceptibility correction	+1.5	0.3
Target-out contribution	−0.4	0.4
Container-wall contribution		0.1
Frequency comparisons		0.02
Root-sum-square of systematic effects		1.4(H_2O)
		1.7($CH_2(CN)_2$)

Several lines of evidence lead to the conclusion that μ^+, H^+, and T^+ (tritons) when slowing down in matter do not reach thermal energy as ions [12]. Below a few hundred eV a positive muon has with high probability permanently captured an electron. Losing energy by molecular collision, it becomes a "hot atom," [13] which, at a few eV, may become part of a molecule, thus retaining its polarization, or may thermalize, probably depolarizing.

A proton in (liquid) H_2O experiences a magnetic field weakened, due to atomic electrons, by 25.6 ppm. When a muon replaces a proton, it should generally experience approximately the same shielding. Fortunately most neutral hydrogen-containing molecules have nearly the same shielding effect as water. We list in table 3 the species expected on the basis of hot-atom work with tritium, also the shift (with respect to protons in water at room temperature) a proton experiences in each. There is a muon-proton difference because the muon, with zero-point energy three times as large, sits higher in its anharmonic potential well, and moves away from its neighbor. We estimate the effect to be about 0.2 ppm in ordinary molecules [14]. However, the muon in a μHO molecule takes part in hydrogen bonding to neighboring molecules, and the higher zero-point energy should lead to a larger hydrogen-bond effect. An estimated upper limit to the additional shielding decrease caused by the hydrogen-bonding effect in water is 4 ppm [15]. We assign 2 ppm for this shift, and an error of ±1.5 ppm.

The results in table 2 for water (combined NaOH and H_2O data) and for $CH_2(CN)_2$ are in gratifying agreement: 1.9 ppm difference, compared with individual errors of 2.8 ppm and 3.1 ppm. We take the average, and, since systematic uncertainties contribute over half the error, we leave the error of the average as 2.8 ppm. The final result is thus $\mu/\mu_p = 3.183347(9)$ (2.8 ppm). (The previously reported results were: Columbia [1], 3.183380(40); Berkeley [2], 3.183369(70); Princeton-Penn [3], 3.183330 (44).)

TABLE 2. *Results for ω_μ/ω_p, including corrections from table* 1
The final ratios include the chemical corrections, and their errors, from Table 3.

Water comparison:	H_2O	3.183350(8) (2.5 ppm)
	NaOH solution	3.183355(8) (2.5 ppm)
	H_2O and NaOH solution combined	3.183350(9) (2.8 ppm)
Final ratios:	$CH_2(CN)_2$	3.183344(10) (3.1 ppm)
Final result:	μ_μ/μ_p	3.183347(9) (2.8 ppm)

TABLE 3. *Muons in water and $CH_2(CN)_2$*
Composition estimated from tritium hot-atom chemistry. δ is the increase in shielding, in ppm, relative to protons in water.

Species	Fraction	δ (proton)	δ (muon)
Water and NaOH solution			
μHO	0.9	0	−2.0
μH	0.1	0.4	0.2
μH_2O^+	≈0	−11	−15
Methylene cyanide			
μH	0.7	0.4	0.2
$\mu HC(CN)_2$	0.3	1.5	1.3
$\mu H_2C(CN)$	<0.1	3.0	2.8

Average shifts: Water, −1.8±2.0 ppm; $CH_2(CN)_2$, +0.5±1.5 ppm.

We now put our results and the recent [16] value of α into the evaluation, by Taylor, Parker, and Langenberg [17], of the muonium hyperfine splitting, ν_m. The predicted value proves to be 4463.289(19) MHz. This is very close to the weighted average of the two most recent results: Ehrlich et al. [4], ν_m = 4463.317(21) MHz; Crane et al. [5], ν_m = 4463.249 (31) MHz. The old discrepancy between hydrogen hfs and muonium hfs, discussed by Ruderman [6] and others, was 40 ppm; it was based on the Columbia muon moment [1] and the 1964 high field muonium [18]. Our result brings the muon moment down 10 ppm; the new muonium results account for the remaining 30 ppm; and the muonum hfs is now in satisfactory agreement with theory using the Josephson-effect α. It is interesting that a more precise value for muonium hfs would lead to a value of α of accuracy comparable to that of the Josephson effect.

Finally, one obtains the muon-electron mass ratio from $g_\mu\mu_e/g_e\mu_\mu$. The result (we follow Taylor et al. [17]) is $m_\mu/m_e=206.7683(6)$.

Acknowledgments

We thank Professor L. Slutsky for much helpful advice on chemical problems, and M. Delay and J. Justice for important contributions. We acknowledge the excellent cooperation of Jimmy Vale and the cyclotron crew, and the valuable contributions of many members of the LRL staff.

References

[1] Hutchinson, D. P., Menes, J., Shapiro, G., and Patlach, A. M., Phys. Rev. **131**, 1351 (1963).
[2] Bingham, G. McD., Nuovo Cimento **27**, 1352 (1963).
[3] Hutchinson, D. P., Larsen, F. L., Schoen, N. C., Sober, D. I., and Kanofsky, A. S., Phys. Rev. Letters **24**, 1254 (1970).
[4] Ehrlich, R. D., Hofer, H., Magnon, A., Stowell, D., Swanson, R. A., and Telegdi, V. L., Phys. Rev. Letters **24**, 513 (1969).
[5] Crane, P., Amato, J. J., Hughes, V. W., Lazarus, D. M., zu Putlitz, G., and Thompson, P. A., Bull. Am. Phys. Soc. **15**, 45 (1970).
[6] Ruderman, M. A., Phys. Rev. Letters **17**, 794 (1966).
[7] A Preliminary report has been given: Hague, J. F., et al., Bull. Am. Phys. Soc. **15**, 608 (1970).
[8] Bailey, J., Bartl, W., von Bochman, G., Brown, R. C. A., Farley, F. J. M., Jostlein, H., Picasso, E., and Williams, R. W., Phys. Letters **28B**, 287 (1968).
[9] This system was used by K. M. Crowe and G. McD. Bingham, ref. [2]. See G. McD. Bingham (Ph.D. Thesis), Lawrence Radiation Laboratory Report UCRL-10107, 1963 (unpublished).
[10] Background was 1 to 1.5% of initial counting rate. Leaving the livetime as an additional free parameter does not change other results, and gives a lifetime within 1 standard deviation of the world average. Our value will be discussed in a future publication.
[11] We are grateful to Prof. L. Slutsky for elucidating the experimental evidence for this statement; the argument will be included in our detailed publication.
[12] Muonium is known to be formed in Ar and Kr. Observation on proton beams shows that as the beam slows down to a few keV it is increasingly neutralized; in various substances the fraction is up to 0.85–0.90 at the lowest energies observable, and still rising. See Allison, S. K., and Garcia-Munoz, M., in Atomic and Molecular Processes, D. R. Bates, Ed. (Academic Press, New York, 1962); Ch. 19. The evidence from hot-atom chemistry is summarized by R. Wolfgang, Prog. Reaction Kinet. **3**, 99 (1965).
[13] See, e.g., Wolfgang, R., op. cit., and Rowland, F. S. J. Am. Chem. Soc. **90**, 4767 (1968). We are indebted to Professor Rowland for helpful advice on hot-atom chemistry.
[14] The first vibrational state of an OH system will have the same stretching as 0μ. From studies of H_2O spectra this is found to be $\approx 1\%$. Ruderman (Ref. 6) finds a rate of change of shielding with covalent bond distance which gives 0.2 ppm for 0.01 A. He also points out the importance of hydrogen bonding; our argument follows his, but with reference to neutral molecules rather than ions.
[15] We are indebted to Prof. J. A. Pople for advice on this question.
[16] Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).
[17] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[18] Cleland, W. E., Bailey, J. M., Eckhause, M., Hughes, V. W., Mobley, R. M., Prepost, R., and Rothberg, J. E., Phys. Rev. Letters **13**, 202 (1964).

DISCUSSION

I. HAMMERMAN: In fitting your decay distribution did you assume that the background was flat?

A. SCHENCK: Yes.

I. HAMMERMAN: Do you have *proof* that it is? For instance, by looking at the time distribution of events in the time interval before capture.

A. SCHENCK: We used different parts of our histogram by using different starting and stopping times for the fit. The results were independent of these times. We looked also at the background without target, and we saw, of course, due to stopped muons, for instance, in the wrapping of the counters, some background. We determined the asymmetry of this background to 0.05, and we measured the frequency of this background too, and we put this all together into our result. It's in there.

Magnetic Moment of the Positive Muon*

D. P. Hutchinson, F. L. Larsen, N. C. Schoen, and D. I. Sober**

Princeton-Pennsylvania Accelerator, Princeton, N. J. 08540

and

A. S. Kanofsky

Department of Physics, Lehigh University, Bethlehem, Pa. 10815,
and
Princeton-Pennsylvania Accelerator, Princeton, N. J. 08540

The magnetic moment of the positive muon has been measured relative to that of the proton by determining its precession frequency in a magnetic field measured by proton NMR. With both protons and muons bound in a spherical sample of water, the result is $\omega_{\mu'}/\omega_{p'}=3.183362\pm0.000030$ (9.4 ppm).

Preliminary measurements of the muon precession frequency in cyclohexane and NaOH aqueous solution yield $[\omega_\mu(\text{cyclohexane})-\omega_\mu(\text{NaOH})]/\omega_{p'}=0.000039\pm0.000031=12.1\pm9.6$ ppm which is reasonably consistent with the difference of -3.6 ppm expected if the muon has the same relative chemical shifts as the proton.

Key words: Muon magnetic moment.

We report a series of measurements of the magnetic moment of the positive muon presently under way at the Princeton-Pennsylvania Accelerator. We have measured the precession frequency of stopped muons in a variety of chemical targets in order to obtain a best value for the ratio μ_μ/μ_p of the magnetic moment of the free muon to that of the proton, and to investigate the possible presence of the chemical correction for the muon in water suggested by Ruderman [1].

The first measurement of the series (Set I), in which only a water target was used, has been reported previously [2]. A longitudinally polarized muon beam was produced by forward decays in flight of a beam of 158 MeV/c pions. The pions and backward muons were stopped in a 1.125-in copper absorber, and the more energetic muons continued through the absorber and stopped in a 2-in-diameter target inside the gap of an 18D40 magnet whose 12.0-kG field was continuously regulated and monitored by proton NMR probes. The field in the target region was mapped with a movable NMR probe containing $NiSO_4$ in water.

The vertical magnetic field causes the muon spin to precess with frequency $\omega=2\mu H/\hbar$ where μ_μ is the magnetic moment. Since the angular distribution of decay positrons is peaked in the muon spin direction, the time distribution of decay positrons at a fixed counter position is modulated by the precession frequency. The ratio ω_μ/ω_p of the muon precession frequency to the proton NMR frequency which measures the magnetic field at the target is equal to the ratio of the muon and proton magnetic moments.

The time interval between the incident muon and the decay positron was recorded by a stroboscopic technique similar to that used in earlier measurements at Columbia [3] which in turn was based on earlier work at Chicago [4]. All decay times were mapped onto the phase interval 0 to 2π by recording only the difference between the length of the time interval and an integral number of cycles of a 163-MHz reference oscillator whose frequency nearly matched the precession frequency of the muons in the magnetic field. Decays were recorded during a 3.95 μs interval which was divided at 1.95 μs into "early" and "late" gates, and the early and late phase distributions were separately stored in 10 channels using a PDP-7 computer. The phase distribution for each group has the form [5]

$$N(\Phi)=\text{const}\{1+a\cos[(\omega_\mu/\omega_0)\Phi-\alpha]\}\exp(-\Phi/\omega_0\tau)$$

where α depends on the detailed geometry and time delays. The difference $\alpha_{\text{late}}-\alpha_{\text{early}}$, however, depends only on the time delay between the early and late gates and the difference between the precession and

* Work supported by the U. S. Atomic Energy Commission under Contract No. AT (30-1)-3217.

** Present address: Meteorologisk Institut, Charlottenlund, Denmark.

reference frequencies. When the frequencies are nearly equal, the expression for the phase difference is

$$\alpha_{\text{late}}-\alpha_{\text{early}}=(\omega_{\text{ref}}-\omega_\mu)[T_1+\langle\Delta t_2\rangle-\langle\Delta t_1\rangle]$$

where T_1 is the early-late delay and $\langle\Delta t_1\rangle$ and $\langle\Delta t_2\rangle$ are constant functions of the gate lengths. This expression is linear in $(\omega_{\text{ref}}-\omega_\mu)$ and passes through 0 when the frequencies are identical.

The initial phases were determined by Fourier analysis of the phase distribution. A linear fit to the phase differences at several proton precession frequencies (magnetic fields) then passes through 0 when the muon precession frequency is equal to the known reference frequency. This occurs at a monitor proton frequency of 51.09974±0.0036 MHz. Corrections were made for the difference between the magnetic field at the monitor probe and the average field in the target (23.0±1.0 ppm) and for the calibration of the clock circuitry. The latter was tested by the use of random pulse distributions modulated by a known RF frequency, yielding an error of −6±4 ppm. Including estimates of other possible systematic errors, the final corrections totaled +17± 6.2 ppm.

In the second experiment (Set II), presently in progress, the method and most of the equipment are the same, but a number of changes have increased the effective rate. The beam line was designed for a larger angular acceptance and a smaller fractional energy spread. The beam momentum was raised to 174 MeV/c and the absorber thickness increased accordingly. In order to increase the counting rate, we have enlarged the homogeneous field region to 4-in square by 2-in high, increased the target size to $3\frac{5}{8}$-in diameter, and enlarged all the counters. The thick counter which timed the decay positron has been replaced by a 6-in wide counter consisting of 1-in wide scintillator segments mounted in a single sir light guide. Within the light guide, each scintillator segment is coupled to a different length of UVT Plexiglas light pipe which delays the light to compensate for the precession of the decay distribution across the counter face. The targets used in the second run thus far are cyclohexane and a 1.0 normal solution of sodium hydroxide in water.

The corrected result for water in Set I, including all systematic as well as statistical errors, is

$$\omega_{\mu'}/\omega_{p'}=3.183362\pm0.000030\ (9.4\ \text{ppm})$$

where the prime indicates a particle in a spherical sample of water. This compares well with the results of previous experiments, 3.183380±0.000040 for Columbia [2] and 3.183360±0.000070 for Berkeley [6].

The earlier experiments assumed that a muon at rest in a water target experiences the same local magnetic field as a proton in water, whose chemical shielding correction relative to vacuum is well known to be 26.3±0.3 ppm [7]. If this is so, then the ratio of free magnetic moments μ_μ/μ_p is equal to the measured frequency ratio in water $\omega_{\mu'}/\omega_{p'}$. Ruderman [1] suggested that the lower mass of the muon may make possible a bond of the form H_2O–μ–H_2O with a shift relative to vacuum of only 5 to 10 ppm.

Since the presence of free OH^- ions may favor the formation of μOH and suppress Ruderman's configuration, one could expect that the addition of NaOH to the water in the target would change the apparent magnetic moment if the Ruderman mechanism were dominant in pure water. A recent precise measurement of the muon magnetic moment in water and several NaOH concentrations by the Washington-Berkeley group [8] has shown no significant difference within errors much smaller than the magnitude of any expected shift.

Since our Set II measurement is still in progress and the final corrections to the data have not been calculated, we are unable to compare our NaOH result to the pure water measurements of Set I and other experimenters. All that we can present here is a comparison between our preliminary results for 1.0 normal NaOH and cyclohexane. The latest uncorrected zero-crossing proton frequencies are 51.10446± 0.00038 MHz (NaOH) and 51.10384±0.00032 MHz (cyclohexane), giving as a difference

$$[\omega_\mu(\text{cyclohexane})-\omega_\mu(\text{NaOH})]/\omega_{p'}$$
$$=0.000039\pm0.000031=12.1\pm9.6\ \text{ppm}$$

Since the proton chemical shielding in cyclohexane is higher than that in water by 3.6 ppm [9], one would expect the shielded muon frequency to be lower than in NaOH solution by this amount if the simple picture of the muon replacing a proton is valid. Our difference of 1.6 standard deviations is consistent with this picture.

Acknowledgements

We are grateful to Prof. M. G. White and the PPA staff for their support of this experiment, and hope that the facilities of the PPA may remain available to future experimenters.

References

[1] Ruderman, M. A., Phys. Rev. Letters **17**, 794 (1966).
[2] Hutchinson, D. P., Larsen, F. L., Schoen, N. C., Sober, D. I., and Kanofsky, A. S., Phys. Rev. Letters **22**, 1254 (1970); also Hutchinson, D. P., et al., in S. Devons, ed., High Energy Physics and Nuclear Structure (Plenum Press, New York, 1970), p. 636.
[3] Hutchinson, D. P., Menes, J., Shapiro, G., and Patlach, A. M., Phys. Rev. **131**, 1351 (1963).
[4] Lundy, R. A., Sens, J. C., Swenson, R. A., Telegdi, V. L., and Yovanovitch, D. D., Phys. Rev. Letters **1**, 38 (1958).
[5] Garwin, R. L., Hutchinson, D. P., Penman, S., and Shapiro, G., Phys. Rev. **118**, 271 (1960).
[6] Bingham, G. McD., Nuovo Cimento **27**, 1352 (1963).
[7] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969); see p. 395.
[8] Hague, J. F., Rothberg, J. E., Schenck, A., Williams, D. L., Williams, R. W., Young, K. K., and Crowe, K. M., Bull. Am. Phys. Soc. **15**, 608 (1970); also these Proceedings.
[9] Pople, J. A., Schneider, W. G., and Bernstein, H. J., High-resolution Nuclear Magnetic Resonance (McGraw-Hill, New York, 1959), p. 89.

Hyperfine Structure Interval of the Ground State of Muonium*

P. A. Thompson, D. Casperson, P. Crane,** T. Crane, P. Egan, V. W. Hughes, G. zu Putlitz,*** and R. Stambaugh

Yale University, Gibbs Laboratory, New Haven, Connecticut 06520

The results of the most recent very weak field determinations of the muonium hyperfine structure are reported. Combined with the older data they give a value of $\Delta\nu = 4463.308 \pm 0.013$ MHz (± 2.9 ppm). The first strong evidence for a nonlinear term in the fractional pressure shift of muonium is presented.

Key words: Argon; krypton; muonium hyperfine structure; muon magnetic moment; nonlinear pressure shift.

Precision measurements of the hyperfine structure of muonium can yield information about the fine structure constant and the muon magnetic moment. Recently the efforts of the Yale Group have been directed to the precision determination of the hyperfine structure interval through measurements of the transition $F=1 \leftrightarrow F=0$ (F is the total angular momentum) in the ground state of muonium at very weak magnetic fields (10 milligauss).

The ground state of the muonium atom is broken into four hyperfine levels ($F=0, 1$; $m_F = \pm 1, 0$). At a field of 10 mG, the transitions (F, m_F): $(1, 1) \leftrightarrow (0, 0)$ and $(1, -1) \leftrightarrow (0, 0)$ are separated by about 30 kHz. The power broadened line width is approximately 500 kHz. Thus these two transitions are excited simultaneously. The resultant line shape is a Lorentzian one.

The experimental technique is similar to that used previously [1]. The magnetic field is achieved through the use of a Moly-permalloy shielding structure. The transition is excited in a high Q cavity and the incoming muons and decay positrons are detected with scintillation counters.

A principal objective of the most recent measurements (August 1970) was to measure the hyperfine structure interval (hfs) by this method at lower gas densities than had been previously achieved by this group. The main problem encountered at low gas densities is deterioration of the signal to noise ratio. This is because the ratio of muons stopped in the active region of the target to those stopped in the walls, counters and other inactive areas decreases with decreasing gas pressure.

In the most recent series of measurements, the diameter of the microwave cavity was increased to 27.666 cm (TM 330 mode). The effect of this was to reduce the number of muons stopped in the side walls of the cavity (see fig. 1). In previous work by this group, it was found feasible to use plastic scintillation counters inside the pressure vessel. In this version, the thickness of the number 5 counter was reduced to 0.5 mm. With 10 atm of argon gas in the target this system gave a muon stopping rate of about 200/s. Of these we determined that approximately 40/s were stopped in the active region of the target. A large portion of the remainder apparently stop in the side walls. Decay positrons are defined as $e_F = 67\bar{3}$ and $e_B = 34\bar{2}$ for the forward and backward directions respectively. Counters 3 and 7 were increased to about 0.7 m in diameter to increase their detection efficiency. The total electron detection efficiency was 20 percent. Numerous other changes including the use of an on-line computer system were made.

As a result of the work of the Nevis Cyclotron Laboratory staff, a new beam was also available for us. This beam had significantly reduced pion contamination compared to those available previously.

With the above changes, it was possible to make a precision measurement of the hfs at 10 atm. The resonance curve obtained in 5 weeks of accelerator time is shown in figure 2. The observed and expected signals are given below. The expected signal is a theoretically calculated signal multiplied by the percentage of muons stopping in the active region.

	Experiment	Expected
S_F	$2.4 \pm 0.2\%$	3.3%
S_B	$1.3 \pm 0.2\%$	2.3%

The differences are presumably due to:

(1) Muon polarization less than 100 percent
(2) Incomplete muonium formation
(3) Pion contamination of the beam.

The presence of the buffer gas perturbs the muonium atom causing a density dependent shift in the observed hfs. Table 1 gives measurements of $\Delta\nu$

* This Research supported in part by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force under AFOSR Grant No. F44620-71-C-0042.

** Present address: Palmer Physical Laboratory, Princeton University, Princeton, New Jersey 08540.

*** I. Physikalisches Institut, Heidelberg, Germany.

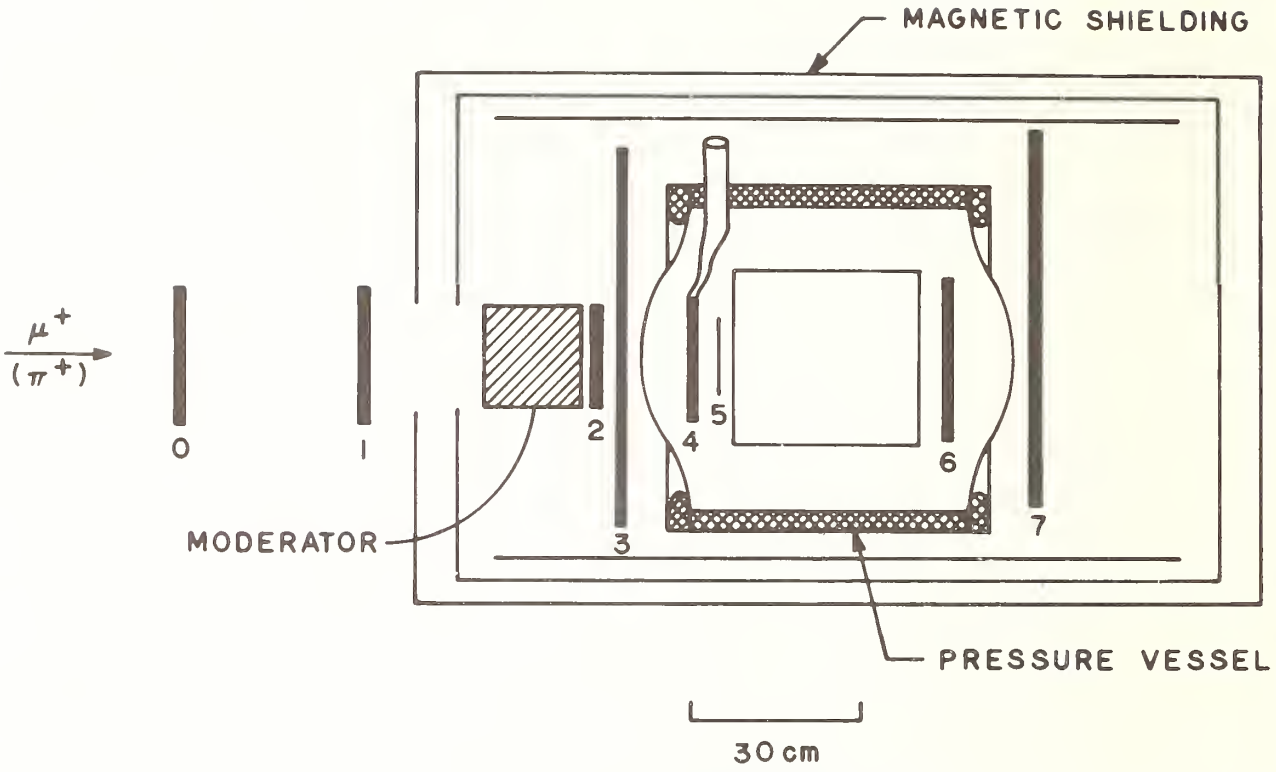


FIGURE 1. *The experimental setup used in the summer of 1970.*
The numbered objects are plastic scintillation counters. Counters 4, 5, and 6 are inside the pressure vessel. A representative light pipe, quartz window, and photo tube are shown for counter number 4; the others are similar.

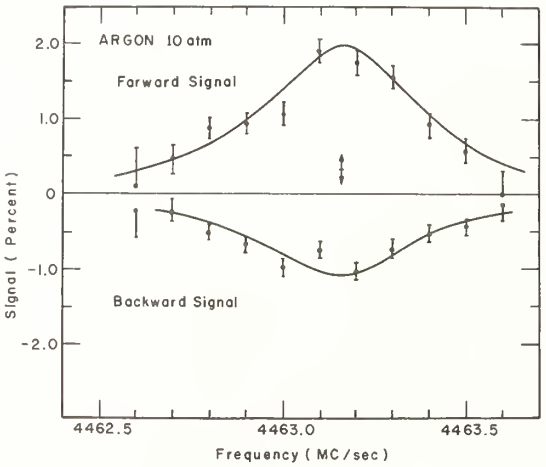


FIGURE 2. *The experimentally observed line shapes at an argon pressure of 10 atmospheres.*
The error bars indicated are statistical.

versus pressure. The results are plotted in figures 3 and 4. One important consideration in extrapolating to a zero pressure value from measurements at finite pressures is whether or not this effect is linear. Recently a quadratic dependence was observed in the pressure dependence of the rubidium hfs in argon buffer gas [2]. Table 2 gives the results of the pressure shift fits. It can be seen that there is strong

TABLE 1. *Measurements of $\Delta\nu$ as a function of gas density*

Argon		Krypton	
Density (psia 0° C)	$\Delta\nu$ (kHz)	Density (psia 0° C)	$\Delta\nu$ (kHz)
137.3	4463 162±13	84.7	4463 066±30
200.9	4463 090±12	142.1	4462 946±19
473.5	4462 778±14	307.0	4462 567±11
948.5	4462 326±16	645.0	4461 755±13
1598.0	4461 718±16	1067.4	4460 809±11

evidence for a non-linear term in argon whereas the case for krypton is ambiguous. This is the first strong (more than 3 standard deviations) evidence for a nonlinear shift in either muonium or any of the hydrogen isotopes. The linear term, $-4.99\pm0.24\times10^{-9}$/torr at 0 °C is in agreement with the value [3], -4.78 ± 0.03, obtained by optical pumping techniques for hydrogen in argon. (The hydrogen value is based on measurements at a fraction of an atmosphere where nonlinear terms are expected to be much smaller.)

Also plotted in figure 3 are two points measured by the Chicago Group at a field of 11 kG, in argon [4]. These points have not been included in the fits

TABLE 2. *Results of pressure shift fits*

		$\Delta\nu(P)=\Delta\nu(0)[1+aP+bP^2]$			
		$a(10^{-9}/\text{torr})$	$b(10^{-15}/\text{torr}^2)$	χ^2/N	$\Delta\nu(0)$ MHz
Argon	Linear ($b=0$)	-4.25 ± 0.10	0	13/3	4463.285 ± 0.010
	Quadratic	-4.95 ± 0.23	7.6 ± 2.6	2.1/2	4463.313 ± 0.015
Krypton	Linear	-9.98 ± 0.08	0	5.2/3	4463.260 ± 0.013
	Quadratic	-10.5 ± 0.50	$+8.7\pm6.6$	3.5/2	4463.296 ± 0.023
	Average				4463.308 ± 0.013

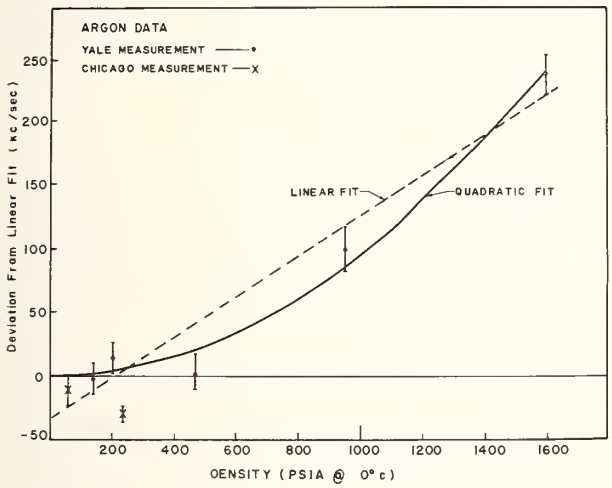


FIGURE 3. *Plot of muonium hyperfine splitting $\Delta\nu$ versus gas density for measurements made in argon.*
To show more detail, a linear portion of the pressure dependence ($a=-4.99\times10^{-9}$/torr @ 0°C) has been subtracted from the directly measured values.

and as can be seen their deviation from our data is appreciable. Interestingly enough the zero pressure intercept of this data is consistent with that of the Yale data.

The errors given are due to counting statistics and are one standard deviation. All known systematic errors are believed to be less than 1 kHz. In particular, the microwave frequency was continuously compared with WWVB to 0.05 kHz; the pressure and temperature were monitored continuously. The gas density was constant to better than $\frac{1}{2}$ psi. (Pressures were measured with a quartz bourdon tube gauge which was calibrated against a dead weight gauge with an uncertainty of 0.1 psi.) The microwave power level was held constant with a fast feed-back system and monitored (stability better than 1%).

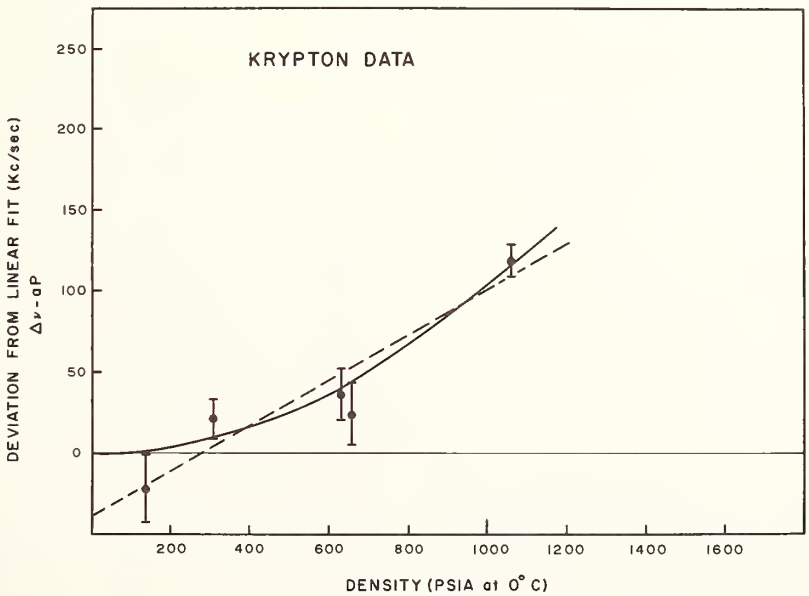


FIGURE 4. *Plot of the muonium hyperfine splitting $\Delta\nu$ versus gas density for measurements made in krypton.*
The same method of plotting is used as in figure 3. The linear term subtracted is $a=-9.98\times10^{-9}$/torr at 0°C.

Taking a weighted average of the results obtained with quadratic fits we get:

$$\Delta\nu = 4463.308 \pm 0.013 \text{ MHz (3 ppm)}.$$

To interpret this result one must consider the theoretical expression for $\Delta\nu$ which is [5]:

$$\Delta\nu = [\tfrac{16}{3}\alpha^2 cR_\infty(\mu_\mu/\mu_{0e})][1+(m_e/m_\mu)]^{-3} \times(1+\tfrac{3}{2}\alpha^2+a_e+\epsilon_1+\epsilon_2+\epsilon_3-\delta_\mu) \quad (1)$$

in which

$$a_e = (\alpha/2\pi) - 0.328(\alpha^2/\pi^2),$$

$$\epsilon_1 = \alpha^2(\ln 2 - \tfrac{5}{2}),$$

$$\epsilon_2 = -(8\alpha^3/3\pi)\ln\alpha(\ln\alpha - \ln 4 + \tfrac{281}{480}),$$

$$\epsilon_3 = (\alpha^3/\pi)(18.4\pm 5),$$

$$\delta_\mu = (3\alpha/\pi)[m_e m_\mu/(m_\mu^2 - m_e^2)]\ln(m_\mu/m_e) - \tfrac{9}{2}\alpha^2\ln\alpha(m_e/m_\mu)[1+m_e/m_\mu]^{-2}.$$

This may be rewritten

$$\Delta\nu = (\alpha^2\mu_\mu/\mu_P) \quad 26\,329\,582 \text{ MHz } (\pm 1 \text{ ppm}). \quad (2)$$

The uncertainty represents the estimated uncertainty due to the higher order terms which have not yet been calculated. At present these are smaller than the experimental limits. If one inserts our experimental measurement and the suggested [6] value of $\alpha^{-1} = 137.03602(21)$ one can obtain a value for the muon magnetic moment:

$$\mu_\mu/\mu_p = 3.183335(14) \text{ 4 ppm}.$$

This may be compared with the other values for the magnetic moment reported in this session; e.g. $\mu_\mu/\mu_p = 3.183347(9)$, 3 ppm, from Hague et al. [7].

Improvements in the apparatus and the availability of much higher intensity meson beams at both Nevis and Los Alamos should allow reduction of the experimental error to less than 1 ppm in the near future. The interrelation of the muon magnetic moment, the hfs and the fine structure constant is apparent in eq (2). One method of determining the muon moment is through measurements of the muonium hyperfine structure at high magnetic field [5]. This is one of the objeetives of the current construction of a two meter high precision solenoid. It seems reasonable that with the envisioned advances it should be possible to test the theoretical expression for $\Delta\nu$ to better than 1 ppm. This is better than the current calculations.

References

[1] Thompson, P. A., Amato, J. J., Crane, P., Hughes, V. W., Mobley, R. M., zu Putlitz, G., and Rothberg, J. E., Phys. Rev. Letters **22**, 163 (1969); Crane, P., Amato, J. J., Hughes, V. W., Lazarus, D. M., zu Putlitz, G., and Thompson, P. A., High Energy Physics and Nuclear Structure, p. 677, edited by Samuel Devons, Plenum Press, NYC (1970).

[2] Ensberg, E. S., and zu Putlitz, G., Phys. Rev. Letters **22**, 1349 (1969).

[3] Brown, R. A., and Pipkin, F. M., Phys. Rev. **174**, 48 (1968).

[4] Ehrlich, R. D., Hofer, H., Magnon, A., Stowell, D., Swanson, R. A., and Telegdi, V. L., Phys. Rev. Letters **23**, 513 (1969).

[5] Hughes, V. W., Ann. Rev. Nucl. Sci. **16**, 445 (1966).

[6] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. of Mod. Physics **41**, 375 (1969).

[7] Hague, J. F., Rothberg, J. E., Schenck, A., Williams, D. L., Williams, R. W., Young, K. K., and Crowe, K. M., these Proceedings.

DISCUSSION

H. Metcalf: You stated that theoretically you don't expect any asymmetries, so I'd like to ask if you in fact looked for asymmetries in your data. Or, put another way, do you fit your data with a symmetrical curve or do you allow for asymmetries in the fit?

P. A. Thompson: We fit our data with a Lorentzian, a symmetrical Lorentzian, and we get extremely good chi squareds. We cannot think of any way to get it asymmetric.

H. Metcalf: Well, have you tried just putting in a small amount of asymmetry to see if chi squared improves?

P. A. Thompson: No, we haven't. It's a good point. It's worth trying. The chi squareds typically come out equal to or less than the number of degrees of freedom. In fact, we've been sort of bothered that they are almost always less than the number of degrees of freedom—sort of an anomalous situation.

H. G. Robinson: I think there are some subtle experimental ways of getting asymmetric lines which have little to do with the physics of the situation. For instance, AM and FM modulation simultaneously imposed on the rf would give you an asymmetric sideband structure, and that would be reflected then in a pulling of the line.

P. A. Thompson: Of course.

H. F. Robinson: So you can't guarantee the absence of it just because you can't think of a theoretical reason for it.

P. A. Thompson: Right. We checked the spectra of the rf power, and it's at the limit of the spectrum analyzer which is a couple of kilocycles wide. The thing that worries us most is the possibility that when you're on one side of the line your microwave power level is higher than on the other side, and we have done everything we know how to say that's less than 1 percent. That's the thing we sit up late at nights worrying about.

W. E. Lamb, Jr.: If you have a Lorentzian which is distorted, then you need really to have a theory of how it's distorted. There is an exception to that. If you stay very close to the center of the Lorentzian, then a linear distortion, a sloping background or something like that, is very much like a shifted Lorentzian, but you can't do much about that unless you know it's there.

E. S. Ensberg: In connection with the quadratic pressure shift, in the absence of any knowledge of what the effect of the three-body collision is, it's interesting to assume that it's simply the same as the binary collisions except that it has the effect of sweeping one of the argons out of the way. The effect of the three-body collision, therefore, is simply to reduce the effect that you predict for the binary collisions, and with reasonable radii for the collision, you get just the numbers that you have for this simple model.

Measurements on atoms of hydrogen
Are not entirely without stratagem.
They give information
And grand confirmation
Of values of alpha that might have been.

R. T. Robiscoe

Determination of the Atomic Hydrogen Fine Structure by Level Crossing in the 2P States of Hydrogen: A Measurement of the Fine Structure Constant *

J. C. Baird,[1] John Brandenberger,[2] Ken-Ichiro Gondaira,[3] and Harold Metcalf[4]

Brown University, Providence, Rhode Island 02912

The Sommerfeld fine structure constant α has been measured by level crossing spectroscopy in atomic hydrogen. The result is $\alpha^{-1}=137.0354(6)$ where the (6) is the combined statistical and experimental uncertainty and represents a 70 percent confidence level. Experimental and theoretical considerations are presented in the light of contributions to systematic errors which might shift the resultant value of α.

Key words; Fine structure; fine structure constant; hydrogen atom; level crossing; α value.

1. Introduction

The fine structure interval $\Delta E(2P_{1/2}-2P_{3/2})$ in atomic hydrogen has been determined using the technique of level crossing spectroscopy. This measurement can be used, along with the well understood theoretical expression for ΔE, to obtain a value for the fine structure constant α. A value of α has recently been determined using a measurement of e/h from the ac Josephson effect in superconducting junctions [1]. Recent measurements of the Lamb shift, S, and $\Delta E-S$ have also been made and result in values for α [1]. In our measurement we observe part of the change of the spatial distribution of resonance fluorescence from the 2P–1S transition when the fine structure levels are crossed by the action of an applied magnetic field. From the value of the crossing field we can calculate ΔE and then calculate the fine structure constant. (While this work was in progress we learned of a similar experiment by Wing, Fontana, and Himmel at Michigan. We have benefited greatly from discussions with Dr. W. Wing.)

* Supported by the National Science Foundation, GP-14104 and the Petroleum Research Fund, PRF 3205-A5.

[1] On leave at the Joint Institute for Laboratory Astrophysics, Boulder, Colo. 80302.
[2] Department of Physics, Lawrence University, Appleton, Wisconsin 54911.
[3] Department of Physics, University of Electrocommunications, Tokyo, Japan.
[4] Department of Physics, State University of New York, Stony Brook 11790

2. Theory

The zero field energy levels for the $n=2$ state of the hydrogen atom, including reduced mass and radiative corrections, are well understood. The precise theory of the Zeeman effect has also been investigated recently and appears to be equally well understood [2]. The effect of degeneracies in resonance fluorescence, or level crossing, has also been discussed and the resulting Breit expression describing the effect appears to be well understood [3]. In this section we summarize the theoretical expressions used in the analysis of our level crossing experiment on the $n=2$ state of atomic hydrogen.

2.1. Zeeman Effect

The Zeeman effect in atomic hydrogen, including the effects of magnetic moments, has been investigated by Brodsky and Parsons and justified from field theory by Brodsky and Primack [2]. [The presently available theory seems to be accurate to a few tenths of a part per million (ppm).] The Zeeman effect is shown, with exaggerated hfs, from a computer plot of the Brodsky-Parsons theory in figure 1.

2.2. Fine Structure Separation

The energy difference ΔE between the $2P_{1/2}$ and $2P_{3/2}$ levels in atomic hydrogen has been summarized by Erickson and Yennie. Also, Barker and Glover

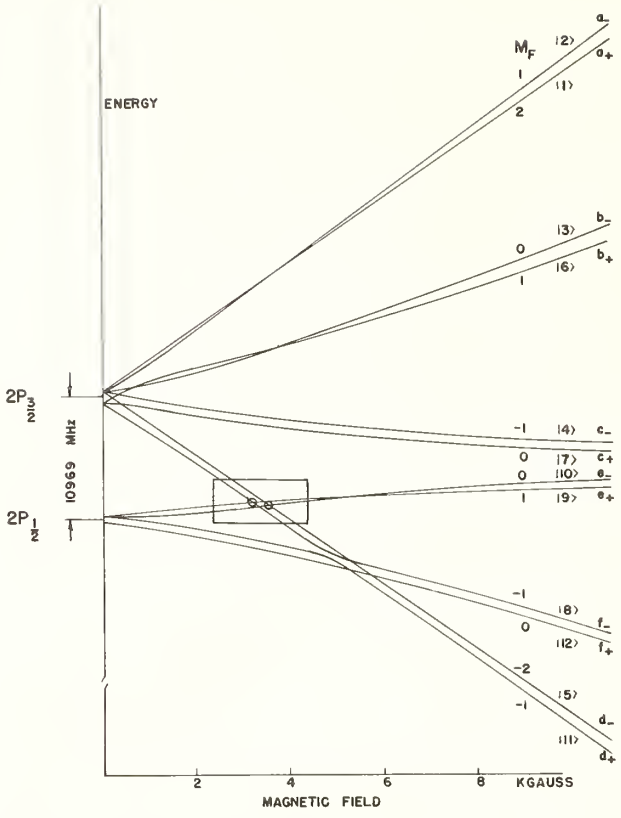


FIGURE 1. *Fine, hyperfine and Zeeman splittings in 2P state of hydrogen for large magnetic fields.*
Circled crossings in box identify fine structure crossings exploited in high-field experiment. The hyperfine structure has been exaggerated and distorted.

and Grotch and Yennie have investigated the reduced mass and/or "recoil" effects. Barker and Glover start from the two particle Bethe-Salpeter equation, reduce it, and calculate matrix elements using non-relativistic wave functions, but containing the reduced mass factor. Grotch and Yennie use an effective potential approach. All authors neglect terms of the order of $\alpha^2 m/M$ and higher. An expression for the fine structure separation for $n=2$ consistent with this is given by [2]

$$\Delta E = \tfrac{1}{16}\alpha^2 Z^4 \mathrm{Ry}_\infty c(\mu/m)^3\left[g_e(m/\mu) - 1 + \tfrac{5}{8}(Z\alpha)^2 + (\alpha/\pi)(Z\alpha)^2(\ln(Z\alpha)^{-2} + \tfrac{1}{3}\gamma + \cdots) + \cdots\right] \quad (1)$$

where $\mu = mM/(m+M)$ is the reduced mass, $\mathrm{Ry}_\infty = e^2/2a_0hc$ is the Rydberg for infinite mass, and $|\gamma| < 1$.

This expression seems to be accurate to better than 0.05 ppm in reduced mass terms. The contribution to the fine structure separation by the smallest radiative term, $(\alpha/\pi)(Z\alpha)^2(\frac{1}{3}\gamma)$, is $\simeq 0.66$ ppm so errors due to radiative effects must be less than this, say 0.3 ppm.

2.3. Level Crossing Theory

The intensity of reasonantly scattered light from an excited multiplet consisting of completely and/or partially degenerate sublevels is described by the Breit formula [3]

$$R = \frac{8\pi^2\rho_1(\nu_1)\rho_2(\nu_2)}{\gamma\hbar^2} \times \sum_{b,b'}\sum_{a_1,a_2} \frac{\langle a_2 \mid H_2 \mid b\rangle\langle b \mid H_1 \mid a_1\rangle \times [\langle a_2 \mid H_2 \mid b'\rangle\langle b' \mid H_1 \mid a_1\rangle]^*}{\frac{1}{2}(\Gamma_{b'}+\Gamma_b) - i(E_{b'}-E_b)}$$

$$= \frac{8\pi^2\rho_1(\nu_1)\rho_2(\nu_2)}{\gamma\hbar^2} \sum_{b,b'} S_{b,b'} \quad (2)$$

and seems to be well established. In eq (2), $(b \mid H_1 \mid a_1)$ is the transition matrix element between a ground state sublevel a_1 and an excited sublevel b, of the operator which annihilates the incident photon with specified direction and polarization. Similarly, $(a_2 \mid H_2 \mid b)$ is the matrix element of the operator which creates the outgoing photon. ρ_1 and ρ_2 are the state densities respectively, of incident and re-emitted light at their average frequencies. Although most of the previous derivations have been based on a semiclassical consideration of the radiation field, treatment of a quantized radiation field leads, of course, to the same result. The main factor $S_{bb'}$ seems to be free from any special assumptions about the temporal and spectral behavior of the incident light. The only important restriction of eq (2) is the assumption of diagonality of the damping matrix, which is valid in our case of Hydrogen 2P states insofar as the mixing with the states of different L or n is neglected.

It should be emphasized that eq (2) describes the total intensity of the emitted light integrated over the frequency: this is not as clearly seen from the semiclassical derivation.

The $b \neq b'$ terms of eq (2) give the coherent scattering intensity or the interference effect between two excited sublevels. The physical meaning of the coherent terms is obvious. The atom may be excited into a mixture of excited states each part of which has a time dependence $\exp(iE_bt/\hbar)$. If $E_b = E_{b'}$ the mixture evolves coherently. If $E_b \neq E_{b'}$ the coherence will be lost in a time interval equal to, roughly, $\hbar/(E_{b'}-E_b)$. If the radiative lifetime $\hbar/\Gamma$ is longer than $\hbar/(E_{b'}-E_b)$, the photon will most probably be emitted after the coherence is lost. No coherent contributions are expected in this case. In fact, the $b \neq b'$ terms have a large denominator so that their contribution is small. In contrast to this, if $\hbar/(E_{b'}-E_b) > \hbar/\Gamma$, the coherence of the states will have more influence on the scattered radiation pattern. The coherence effect manifests itself as a change in the re-emitted light intensity in a specified direction.

For the $|\mu - \mu'| = 2$ crossings (μ = magnetic quantum number), with which the present experiment is concerned, the expression for the line shape, including the dependence of the signal on experimental geometry, can be obtained easily from the general expression [4].

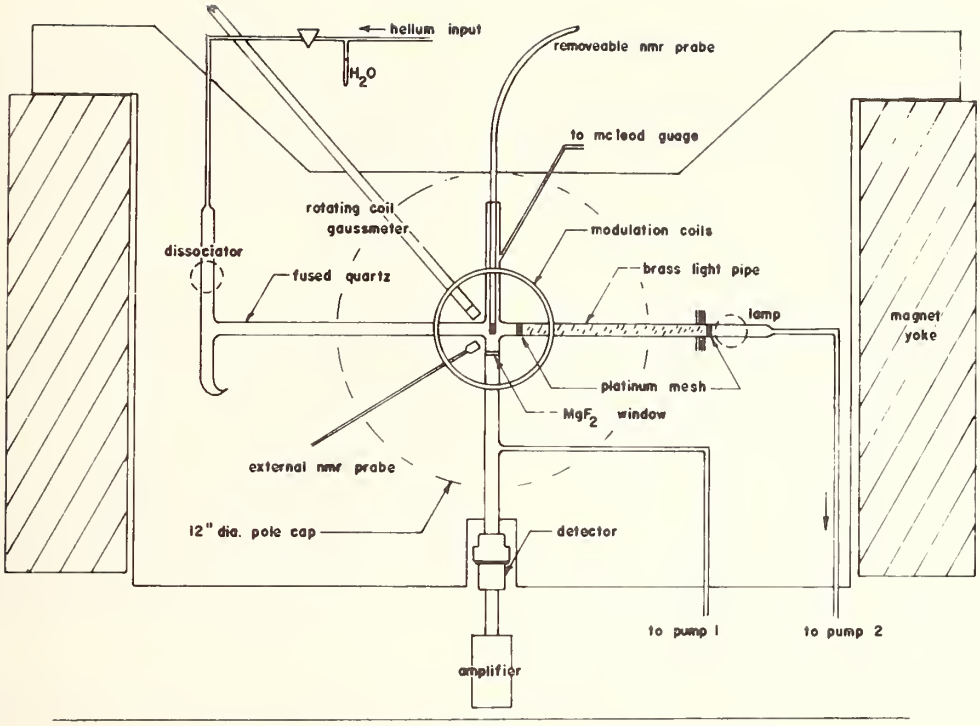


FIGURE 2. *Diagrammatic view of apparatus in 12 in magnet.*
Helium and hydrogen atoms flow from left to right. Magnetic field into the page.

We limit our discussions to the 3484 gauss crossings indicated by the box in figure 1. There are four level crossings in this region. One of them has $|\mu-\mu'|=3$, which precludes simultaneous excitation and hence does not result in a crossing signal. The signal from the $|\mu-\mu'|=1$ crossing turns out to be almost forbidden according to the Δm_I rule. Moreover, if one adopts $\theta_1=\theta_2=\pi/2$ geometry the $|\mu-\mu'|=1$ signal vanishes exactly. There remain two $|\mu-\mu'|=2$ crossings. If in addition to $\theta_1=\theta_2=\pi/2$ we chose $\phi_1-\phi_2=\pi/2$, the signal becomes a pure Lorentzian, but if $\phi_1-\phi_2$ deviates from $\pi/2$, the signal will contain a small dispersion type component.

Although the qualitative features of the signal should be clear from the above discussion, we must actually diagonalize the total Hamiltonian numerically and compute the energy eigenvalues and the mixing coefficients in order to determine a precise theoretical line shape. We have used the F, m_F representation, because of the availability of the Brodsky and Parsons [2] matrix elements. Each of the two hfs coherent contributions to the line shape has a $\simeq$0.1 percent asymmetric shape due to the field dependence of the mixing coefficients, and the two components differ in strength by about 0.05 percent, due to inequivalent mixing of hyperfine states. The ratio of the coherent to incoherent contributions is 0.138 at the peak position, which is to be compared with 0.146 if the hyperfine interaction is neglected.

From eq (2) we may derive an expression, $S(H)$, for the level crossing lineshape having a width given by $\Gamma_{bb'}=(\Gamma_{b'}+\Gamma_b)/2$ and whose field dependence is that for the Zeeman effect in atomic hydrogen.

3. Apparatus

The apparatus consists of a hydrogen atom source, a lamp, optical system, detector, amplifiers, magnetic field modulation and ancillary components mounted rigidly as a unit and inserted between the pole pieces of an electromagnet, figure 2. Commercial tank helium, containing water vapor and molecular hydrogen impurities, flows through a 2450 MHz dissociator discharge which produces hydrogen atoms. The hydrogen atoms are carried by the helium into the scattering region where they absorb and re-emit resonance radiation from the lamp. The atoms continue beyond the scattering region where they recombine on a platinum mesh or on the walls of a brass light pipe, and flow on as molecules through the light pipe into a second microwave discharge. The absence of atoms in the light pipe reduces self reversal. The second 2450 MHz discharge forms the lamp where hydrogen molecules are dissociated and excited to produce Lyman-α (1216 Å) radiation. The helium carrier gas and products of the discharge are exhausted through the rear of the lamp by a mechanical pump. Radiation scattered by atoms in the scattering region is detected by a nitric oxide photoionization detector mounted in a direction perpendicular to both the incident light and the external field.

3.1. Hydrogen Atom Source

A 2450 MHz microwave discharge dissociates hydrogen atoms from wetted tank helium flowing through 13 mm O.D. quartz tubing. The dissociation efficiency is stabilized by rigid mounting of the dissociator cavity and stabilization of the discharge temperature by careful regulation of the dissociator cooling air.

The atom source is about 10 cm outside the 12 in magnet pole caps and is placed so that light from the dissociator will not reach the scattering region. The H-atom concentration reaching the scattering region is estimated to be about 10^9 atoms/cm^3. 2450 MHz was chosen to reduce stray a-c fields by the use of cavities and microwave cutoff of wave guide attached to the cavities.

3.2. Lamp

The lamp consists of a 5 cm section of 13 mm O.D. fused quartz tubing terminated at the front by a grounded 0.010 in platinum mesh which confines the discharge and reduces self reversal. The quartz tubing, platinum mesh, and brass light pipe butt together in a single epoxy joint cooled by a brass radiator attached to the light pipe. The discharge is excited by an air-cooled 2450 MHz microwave cavity. Lamp noise is reduced by rigid mounting of the excitation cavity and by carefully regulating cooling air. The lamp discharge is located about 21 cm from the scattering region and about 8 cm from the edge of the 12 in magnet pole caps; this reduces the solid angle and the field dependence of the lamp at the expense of intensity at the scattering region. The lamp output is estimated at 10^{11} photons/s. The lamp's field dependence varies with helium pressure and is less than a few percent change over several hundred gauss at the scattering region.

3.3. Optical System

The front of the scattering region and of the lamp are defined by platinum gauzes used in place of MgF_2 windows which were found to deteriorate under direct illumination. The transmission of the gauze is a constant 80 to 90 percent in contrast to the ever decreasing MgF_2 transmission which originally never exceeds 55 percent at Lyman-α.

The light path from lamp to scattering region consists of a brass light pipe which helps provide electrical ground for the platinum gauzes as well as provide shielding of the scattering region and NMR probes from microwave fields by means of microwave cutoff.

Scattered light proceeds through an MgF_2 window separating the scattering region from an evacuated light pipe attached to the photoionization detector. There are no MgF_2 degradation problems here since this window only passes scattered light whose intensity is greatly reduced relative to the incident radiation.

3.4. Detector

The Lyman-α radiation scattered at 90° passes through the MgF_2 window and an evacuated light pipe and falls on the detector which is a nonmagnetic nitric oxide photoionization chamber. Its response is limited to the 1150–1350 Å region by its window and filling gas.

3.5. Amplifiers

The amplifier chain in the apparatus consists of preamplifiers at the photodetector and at the HR-8 lock-in detector. Because the photocurrent is quite small, typically 50 pA, the detector is mounted directly on a dc-coupled FET follower impedance matching amplifier with an input impedance of 10^9 Ω. The dc output of the amplifier drives one channel of a Hewlett Packard 7100 recorder while the ac output drives a PAR Type A preamplifier to the lock-in amplifier. The output of the HR-8 is applied to the other channel of the recorder, and hence both the dc and ac signals are recorded on the same chart.

3.6. Magnet

The magnet is a 12 in Harvey-Wells electromagnet with a 2.25 in gap. Its homogeneity over the 20 cm^3 that includes the scattering region as well as the external NMR probe is about 10 ppm. The magnet supply stability is 10 ppm short term (~0.05 s) and 5 ppm long term (~10 s).

3.7. Field Modulation

The HR-8 lock-in supplies a sinusoidal reference signal which is amplified and transformer coupled to magnetic field modulation coils. The modulation frequency is about 17 Hz and the amplitude is about 23 G p–p. A transformer coupling guarantees a vanishing dc component of the modulation.

3.8. Data Acquisition and Processing

The outputs of both the lock-in detector (signal) and the NMR frequency (field) are converted to digital form and a nonlinear least squares computer program fits a theoretical line shape to the data. From this procedure we obtain the NMR frequency corresponding to the average location of the hfs level crossings and correct this by 11 ppm to find the fine structure separation.

The proton resonance frequency is measured with a Hewlett Packard[5] 5245L frequency counter which is interfaced with a paper tape punch through a parallel-to-serial converter built from Digital Equipment Co. logic blocks. The counter is also equipped with a 5265A DVM plug-in, which monitors the output of the lock-in amplifier, figure 3. By switching the function of the counter, frequency or voltage measurements are made and recorded digitally on

[5] Commercial equipment identified in this paper in order to accurately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Bureau of Standards.

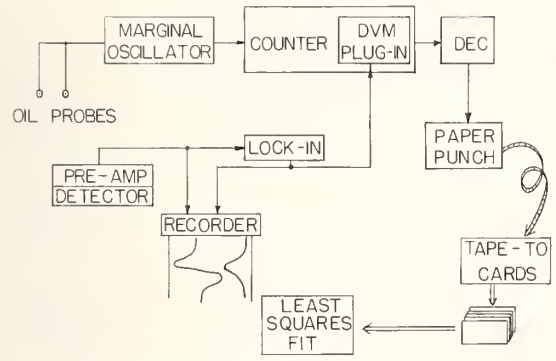


FIGURE 3. *Data acquisition and processing.*

the paper tape. Each data point consists of 10 to 20 consecutive measurements of the signal strength (voltage) with a frequency measurement at both the beginning and the end of the sequence. The average of the two frequencies and the average of all the voltages then constitute a data point.

3.9. NMR Probe for Field Measurement

Highly accurate d-c magnetic field measurements are necessary since the fractional uncertainty in the field measurement contributes a fractional uncertainty in α^{-1} half as large. We use a marginal oscillator equipped with a dual probe designed such that the field at one of the oil samples can be biased slightly. Then any field difference that exists between the two probes can be reduced to less than 1 ppm very easily by merely varying the bias, and most important, this can be done in spite of 5 or 10 ppm fluctuations in the main field.

The external probe consists of an approximately spherical 5 mm diam sample of mineral oil centered in a 7 mm diam, 5 cm long solenoid whose axis is oriented parallel to the field of the magnet (fig. 4).

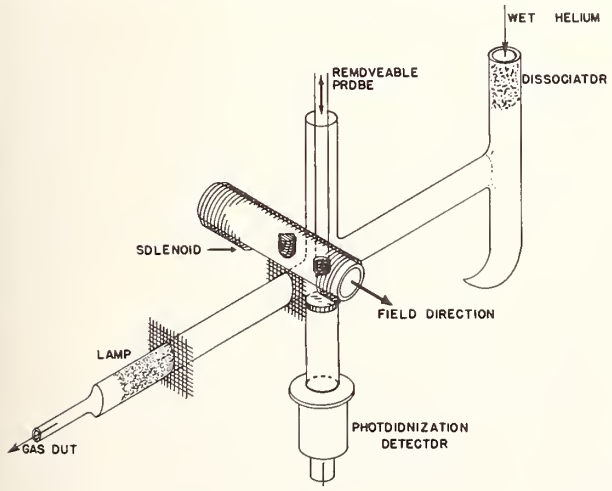


FIGURE 4. *Orientation and position of biasing solenoid that contains the external NMR probe.*
Diagram also depicts removable internal probe, lamp, dissociator, MgF_2 window, and photoionization detector.

FIGURE 5. *Superposition of modulation bucked NMR, narrow signal, and NMR in the modulation field used for the level crossing, broad signal.*

A solenoid configuration was chosen to bias the external probe because its small flux leakage facilitates exclusive biasing of the external probe with minimal degrading of the homogeneity in the scattering region even though the separation between the centers of the solenoid and scattering region is only 2 cm. The second probe is an internal probe that can be moved in and out of the scattering region for the purposes of setting the bias current. Once the bias current is set, the internal probe is withdrawn.

The bias current is determined by optimizing the superposition of the two NMR resonances, one from each probe, displayed on a scope. Using a 17 Hz modulation of several gauss amplitude, the ringing resonances were easily superimposed since a slight deviation from exact superposition produced interference effects readily observable on the scope.

Another feature of the field measuring system concerns the amplitude of the proton resonance modulation. The 23 G p–p amplitude necessary for the level crossing signal was far too large for accurate NMR measurements; therefore, a second solenoid was wound directly over the d-c bias solenoid. The dimensions and wire size were chosen so that when wired in series with the Helmholtz modulation coils, the modulation field inside the solenoid was reduced to about 1/13 the value outside. Figure 5 shows a lock-in amplifier trace of the NMR signal from this configuration of probes and modulation fields.

4. Experimental Data

All data are processed by a computer program that calculates a theoretical a-c line shape $S(H)$ characterized by six adjustable parameters, $V(I)$, $I=1$ to 6,

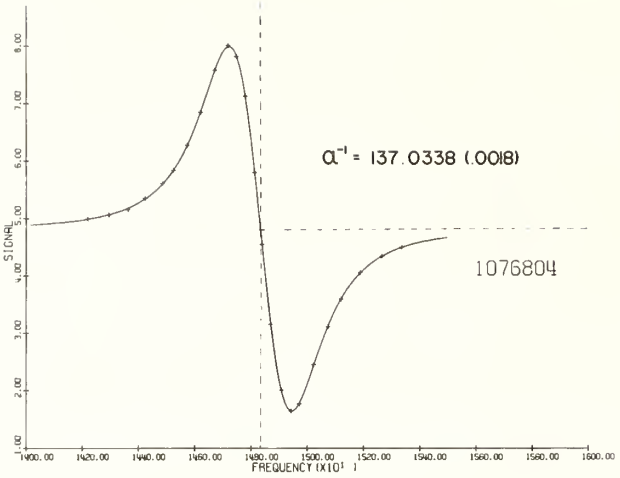


FIGURE 6. *Experimental points fitted to theoretical lineshape for experiment 1076804.*

and fits it to the 25 or so data points. The fitting program centers around an unweighted, nonlinear least square sroutine that varies the V(I) until the sum of the squares of the residuals between the theoretical line shape and experimental data points is a minimum. This routine works as follows: starting with initial estimates of the V(I), a theoretical line shape $S(H)$ and its partial derivatives are evaluated at the 25 or so data points (field points). Products and cross products of the derivatives are summed and arranged to form the normal matrix which describes the overall sensitivity of $S(H)$ to changes in the V(I). The inverse of the normal matrix, the partial derivatives, and the residuals are used to calculate an improved set of parameters V(I). When the process is iterated several times, the sum of the squares of the residuals converges to a minimum.

4.1. Working Line Shape

The theoretical line shape can be calculated accurately from the Breit formula as discussed in section 2. If we neglect the field dependence of the matrix elements and account for the curvature of the Zeeman levels in an approximate way, the line shape can be written

$$S(f)=A\left(\frac{1+\beta w z_+}{1+(wz_+)^2}+\frac{1+\beta w z_-}{1+(wz_-)^2}\right)$$

$$z_\pm=[f-f_0\pm\Delta][1-aw(f-f_0\pm\Delta)] \qquad (3)$$

where f is the field variable measured in proton resonance frequency units, A is the signal amplitude, w is the inverse of the width ($\Gamma_{b'b}$) of the signal, in units of (proton resonance frequency)$^{-1}$, β is the dimensionless asymmetry parameter, f_0 is the position of the center of the level crossing signal in units of proton resonance frequency, and Δ is $\frac{1}{2}$ the hyperfine splitting in the same units. The factor $a=0.0002328$ is a correction factor for the curvature of the Zeeman levels. This line shape involves several approximations which have been discussed in section 2. These approximations involve neglect of a slight asymmetry in the two coherent line shapes and a field dependence of the incoherent scattered light. These approximations have negligible effect in determining α^{-1}.

The approximate line shape, or working line shape is treated by a subroutine which simulates the effect of modulation and lock-in amplification to high precision. The least-squares program then varies f_0, β, A, w, Δ, and the baseline (offset) in order to determine the best fit to the data. An example of a data fit is shown in figure 6.

4.2. Fine Structure Interval

The frequency f_0 is directly related to the value of the fine structure separation, and therefore α^{-1}, and is the most important parameter. By diagonalizing the Brodsky-Parsons Hamiltonian submatrices as a function of field and using the fine structure formula we find $f_0=(16\,690.1844\alpha)^2$ where the theoretical f_0 is expressed in KHz, and was chosen on the basis of the program tests mentioned in section 5.

The field dependent terms in the Brodsky-Parsons matrix elements were calculated using $\mu_B H=328.7338$ hf for a spherical sample of oil. This result has been derived from the 3.7 ppm chemical shift between H_2O and Nujol, the experimental value of g_e, and the expression $g_j(\mathrm{H}\ ^2S_{1/2})=g_e(1-\frac{1}{3}\alpha^2)$.

Data were taken at helium pressures ranging from 0.635 torr to 5.2 torr and at argon pressures around 1 torr. Nine runs were made using argon as the carrier gas; these data are statistically indistinguishable from the helium data insofar as values for the fine structure constant are concerned.

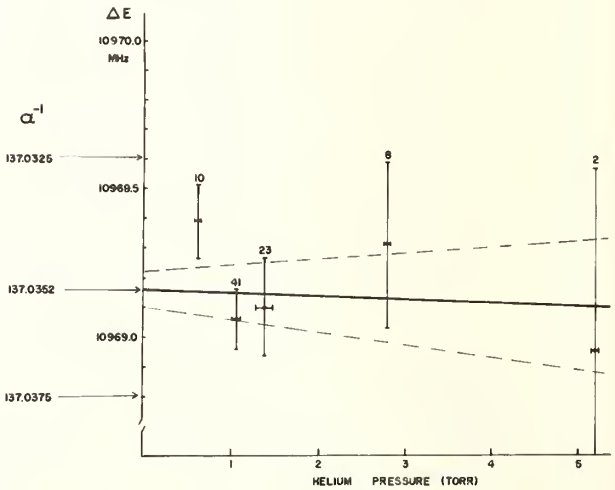


FIGURE 7. *Hydrogen fine structure splitting* ΔE *as a function of helium pressure.*

Error bars indicate one standard deviation in the scatter of the measurements of ΔE at each pressure. Solid line is a linear least-squares fit to the five data points which are weighted in proportion to the number above the error bar. Dashed lines differ from the solid line by one standard deviation in both slope and intercept.

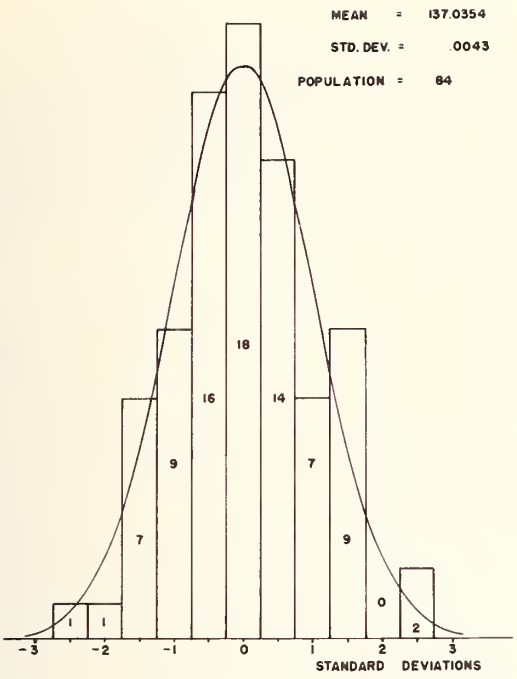


FIGURE 8. *Distribution of values of α^{-1} for 84 runs.*

The measured $2P_{3/2}$–$2P_{1/2}$ splitting ΔE is shown in figure 7 as a function of helium pressure. The solid line is a linear least squares fit to the data points. If we assume there is negligible pressure shift and average all the data we obtain $\alpha^{-1}=137.0354$ and the standard deviation is 0.0043. The statistical distribution of the values of α^{-1} is shown in figure 8.

This includes an additional $+0.0001$ correction for the Stark shifts arising from motional electric fields. This mean value of α^{-1} compares favorably with the intercept $\alpha^{-1}=137.0352$ of figure 7.

4.3. Line Width

As a part of the fit of the experimental data we obtain the line width, or the damping coefficient $\Gamma(P)=1/\tau$. The slope of the graph of $\Gamma(P)$ vs P, 0.0695 (ns-torr)$^{-1}$, is used to calculate the Lorentz broadening cross section for hydrogen Lyman-α in helium, $\sigma_L=95(5)\times10^{-16}$ cm^2. Accurate measurements of Γ_0 are frustrated by the effect of coherence narrowing which is most prevalent at low helium pressures in the present experiment.

5. Systematic Errors

5.1. Working Equation

The expression for the line shape derived from the general level crossing equation includes several approximations which lead to an offset in the value of α^{-1} of less than 0.1 ppm. The values for the fine structure interval are shifted by this amount when the computer calculates α^{-1} from the data (we consider the effect to be insignificant).

The dependability of the fitting program and the working equation was checked and the fit of the two line shapes was excellent; the rms deviation between the test and fitted line shapes was only 0.002 percent (20 ppm) of the signal amplitude peak-to-peak.

By diagonalizing the Brodsky-Parsons Hamiltonian submatrices (using appropriate fundamental constants) as a function of field near the crossing points one obtains $f_0=(16\,690.1861\alpha)^2$ where f_0 is the average of the hfs crossing points measured in NMR frequency units in kHz. The expression we actually use is $f_0=(16\,690.1804\alpha)^2$ and was chosen on the basis of the program tests described above, the criterion being that α^{-1} put into the test data should agree with the α^{-1} obtained by the fitting program. The 0.1 ppm difference represents the *net* correction that we make to the energy level expression in order to account for line shape and line shape approximation effects.

We also find that the mean value of α^{-1} determined from the computer fit of the data is insensitive to arbitrary changes by factors of 2 in the sizes of w, Δ and the modulation amplitude when these are used as fixed input data. This is another way of stating that there is low correlation between f_0 and these parameters in fitting the line shape as can be seen by examining the variance-covariance matrix.

5.2. Dispersion Component in the Level Crossing Lineshape

The observed signal is composed of a Lorentzian part and a dispersive part. The asymmetry due to the dispersive part of the signal causes a shift in the signal maximum and would give rise to an apparent change in the value of f_0, the crossing field, and therefore α^{-1} were it not for the fact that the computer fit to the data interprets the lineshape correctly and determines the amount of this dispersive asymmetry. If the systematic errors all arose from the dispersion component, then the program correctly unravels the lineshape and there are no unusual problems.

Any deviation from orthogonal scattering geometry can produce a dispersion-like asymmetry. Since the detector and lamp both subtend a finite solid angle there is ample opportunity for such asymmetries to appear in our signal. The maximum possible value of the dispersive signal due to purely geometrical effects is $|\beta|\leqq 0.05$, before averaging over the cylindrical symmetry and limited by dimensions of the optical system. The average of 84 runs yields $|\bar{\beta}|=0.022$ and $\bar{\beta}=0.016$.

We may estimate the expected value of β due to the optical depth of the scattering atoms. We find $\langle\beta\rangle\simeq 0.5$ percent which is within a factor of 2 of our average observed value of β. Considering the complexity of the physical situation, this is a reasonable agreement.

We have established that the optical depth asymmetry effect displays the properties of a geometrical asymmetry under field reversal, can be roughly calculated from reasonable assumptions about the hydrogen atom concentration, and behaves in a predictable manner when the hydrogen atom concentration is

varied. The bulk of the contribution to β stems from the hydrogen atom concentration, and in particular, from the optical depth effect.

5.3. Modulation

In anticipation of our computer lineshape analysis, and with the possibility of some slight distortion of the 17 Hz carrier with its sidebands, the purity of the modulation waveform was investigated. The center of a 20 Hz passband was swept from 15 Hz to 150 Hz, and at typical modulation levels, all harmonics of 17 Hz were more than 60 dB below the fundamental. The largest source of distortion was a 60 Hz component that was down by 50 dB. We therefore believe that lineshape distortions from these sources are not important.

5.4. Electronics

The insensitivity of the lock-in amplifier gain to signal level at constant phase (gain stability) was found to be 0.1 percent; the lock-in stability under 180° phase reversal of a constant input signal was measured to be better than 0.05 percent. The gain stability-quantum efficiency product for the photoionization detector was found to be linear to better than 1 percent for light levels and accelerating voltages used.

Odd-amplifier distortions considered all together lead to 1.44 ppm shifts in α^{-1}. The largest distortion is due to the photoionization detector. This is an even-amplifier distortion since it depends only on the light intensity and not on the phase of the signal at the crossing point, f_0, leading to a possible shift in α^{-1} of 0.44 ppm. The mixer in the lock-in amplifier may have a gain which depends on the signal phase contributing 0.74 ppm shifts to α^{-1}.

If we take each of these effects as contributing to a systematic error we have 1.7 ppm.

5.5. Lamp Profile and Field Dependent Background Light

The several lamp profile effects that can produce distortions in the level-crossing signal are the Doppler lineshape, lamp self-reversal, and variations of the lamp intensity with magnetic field.

The most serious effect concerns the direct interaction of the lamp discharge with the magnetic field. This leads to a field dependence of the lamp intensity that is more or less linear with field. The average slope of this lamp-variation is 1.5 percent per 250 G.

A lamp variation correction $[1+b(f-f_0)]$ is applied directly to the experimental data prior to fitting, b having been determined with an accuracy of better than 10 percent from the d-c level crossing signal. Since the correction is typically 8 ppm, there may be an error as large as 1 ppm. The d-c signal is monitored for more than 10 linewidths in order to determine the value of b as accurately as possible and in order to ascertain that the linear lamp variation is a sufficiently accurate description of the lamp field dependence. This is done for each run.

The effect of the emission profile on the center of the level crossing signal is very small since the Doppler width is about 10 times larger than ΔE and 1000 times larger than the width of the level crossing signal. The fractional change of light intensity is 0.001 over the width of the signal. This effect appears as a field dependence of the scattered light and is corrected by the normalization of the data before the lineshape is fitted to it.

5.6. Field Measurement

During the accumulation of data the NMR signal from the external probe was displayed on an oscilloscope. The NMR frequency for which the resonance pattern was centered symmetrically on the oscilloscope was taken as the field measurement. Since the width of the oscilloscope trace corresponded to 1.7 G, individual measurements were accurate to about 1/50th oscilloscope width or $1.7/50=35$ mG (10 ppm). A careful investigation was made to determine how much of this 10 ppm uncertainty was random.

By switching the modulation amplitude between 1.7 G and much smaller values it was established that any systematic error associated with the oscilloscope method of frequency determination was less than 1 ppm.

The external NMR probe is calibrated by biasing its field to equal the field at the center of the scattering cell. Ability to reset the bias current was about ± 3 mG or less than $\frac{1}{2}$ ppm in a α^{-1}.

The rf spectrum of the marginal oscillator was observed to be pure, regardless of whether the oscillator was on or off resonance. This was done by heterodyning with a BC 221 frequency meter.

The diamagnetic correction due to the slight nonspherical geometry of the NMR probe is insignificantly small.

5.7. Stark Shifts

The $2s$ and $2p$ states of hydrogen are uncoupled in a pure magnetic field but mixed in the presence of an electric field. The mixing leads to unequal shifts of the levels important to the level crossing experiment and thus a shift of the crossing frequency f_0. The magnitude of the effect is approximately $\Delta\alpha^{-1}=0.2E$ ppm where E is in volts/cm.

The relative shift by a motional electric field, which is certainly unavoidable since an applied strong magnetic field is one of the essentials of the experiment is approximately $0.5\times10^{-8}T$. Here T is temperature in K in the scattering cell which was measured to be room temperature. The relative shift of the crossing magnetic field therefore leads to a correction of $+0.75$ ppm $=+0.0001$ to α^{-1}. This correction is applied to the values of α^{-1} obtained from the fitting program.

5.8. Light Shift

Strong, incident light can shift the energy levels causing an error in the determination of α^{-1}. The relative shift to the fine structure splitting is 0.92×10^{-8} for typical light levels and therefore we conclude that the light shift can be completely neglected.

5.9. Population of HFS Levels

We have assumed that the hfs levels of the ground state are equally populated when deriving the expression for the level crossing signal. At thermal equilibrium there will be a Boltzmann distribution with $\Delta N/N \sim 2\times10^{-7}$ for the difference in population between the ground state levels taking part in the level crossing. The ground state atoms are assumed to be in thermal equilibrium since they must travel 21 cm down a quartz tubing, in the presence of less than 5 torr of He, before reaching the scattering region.

5.10. Pressure Dependence of ΔE

We have assumed that there is virtually no pressure dependence on ΔE, and have consequently averaged all data. We believe the effect of pressure on the fine structure separation to be small since the fine structure interval is between two states differing only in J. If we assume the cross section for a J-dependent level shift to be as large as 10^{-14} cm^2 we obtain, using an order of magnitude calculation, $\Delta E/E \simeq 10^{-7}$ for the largest pressures used. Several runs were made at differing pressures, but no statistically significant pressure effect could be ascertained.

6. Conclusions

We now summarize all the above estimates of possible systematic errors, in terms of their effect on the measurement of α^{-1}:

1.5 ppm	magnetic field measurement
1.0 ppm	field dependence of lamp intensity
1.7 ppm	non-linearity in electronics
0.1 ppm	motional Stark effect
0.1 ppm	treatment of lineshape asymmetry
0.2 ppm	uncertainty in natural constants [5]
0.5 ppm	uncertainty in theory.

Combining these vectorially, we get $\sigma_2 = 2.5$ ppm as a compounded estimate of the probable systematic uncertainty in our value of α^{-1}.

6.1. Final Result for α^{-1}

By combining $\sigma_1 = 0.00047$ and $\sigma_2 = 0.00034$ we obtain the total uncertainty $\sigma = 0.0006$ associated with the present measurement of the fine structure constant. Thus our final result is:

$$\alpha^{-1} = 137.0354(6)$$

where the (6) represents a 70 percent confidence limit. This corresponds to a $2P_{1/2}$–$2P_{3/2}$ fine structure interval of 10 969.13(0.10) MHz.

It is our belief that the signal to noise can be increased in our experiment and that the error can be reduced to approximately $\frac{1}{2}$ the present value.

7. Acknowledgements

Numerous physicists have given advice to us. We would especially like to thank our colleagues at Brown, members of the atomic physics groups at Yale, Harvard, M.I.T., and Michigan.

8. References

[1] Most of these measurements are reported elsewhere in this volume.
[2] Brodsky, S. J., and Parsons, R. G., Phys. Rev. **163**, 134 (1967); Brodsky, S. J., and Primack, J., Phys. Rev. **174**, 2071 (1968).
[3] Weisskopf, V., Ann. Phys. **9**, 23 (1931); Breit, G., Rev. Mod. Phys. **4**, 504 (1931); Breit, G., Rev. Mod. Phys. **5**, 91 (1933); Goertzel, G., Phys. Rev. **70**, 897 (1946).
[4] Rose, M. E., and Carovillano, R. L., Phys. Rev. **122**, 1185 (1961); Kretzen, H., and Walther, H., Physics Letters **27A**, 718 (1968).
[5] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

DISCUSSION

M. Leventhal: How much dispersion did the computer ask for?

H. Metcalf: Typically, the dispersons were in the range of 2 to 3 percent of the total amplitude. We believe that the dispersion arose from the optical depth of the scattering gas, and so we tried to vary that to see if the computer would ask for a larger dispersion. And in fact, it asked for a larger dispersion, and of the size that we had expected. Also, at the same time we observed a substantial increase of coherence narrowing. We think that this assumption that the dispersion arises from the optical depth of the scattering atoms is therefore valid. The range of these dispersion values was typically over a factor of 4, perhaps from a half a percent to 2 percent.

N. F. Ramsey: What do you feel are the particular advantages of the level crossing as opposed to maybe bringing the levels reasonably close together and then using an oscillatory field?

H. Metcalf: For one thing, one doesn't have to worry about rf power level. In addition, if one is going to bring the levels reasonably close together, it seems to me that one should just continue and go all the way.

N. F. Ramsey: Well, it's not very difficult to—

H. Metcalf: I don't see any reason why both experiments can't be done.

N. F. Ramsey: Sure.

H. Metcalf: Of course, in this particular case the transition is kind of weak. It's not an electric dipole transition. The strong rf fields required would produce rather large stark shifts.

A Measurement of the $2^2S_{1/2}$–$2^2P_{3/2}$ Energy Separation, $(\Delta E\text{-}\$)$, in Hydrogen $(n=2)$*

Tong Wha Shyn, R. T. Robiscoe,*** and W. L. Williams**

Randall Laboratory of Physics, University of Michigan, Ann Arbor, Mich. 48102

We have measured the $2S_{1/2}$–$2P_{3/2}$ energy separation, $(\Delta E-\$)$, in the $n=2$ state of atomic hydrogen by an atomic beam radiofrequency method. We used a beam of metastable atoms in the single hyperfine state $\beta_A(F, m_F=0, 0)$. The electric dipole transitions $\beta_A-b(F, m_F=1, +1)$ and $\beta_A-d(F, m_F=1, -1)$ were observed at fixed frequencies by sweeping the Zeeman magnetic field parallel to the beam axis. Our final value of the $(\Delta E-\$)$ is 9911.250 ± 0.063 MHz where the quoted uncertainty is one average deviation from the mean. Using the revised experimental value of Lamb shift, $\$_H=1057.90\pm0.10$ MHz, we obtain $\Delta E=(10969.15+0.12)$ MHz. The corresponding value for α^{-1} is 137.0356 ± 0.0007.

Key words: Energy separation $(\Delta E-\$)$ in $H(n=2)$ and fine structure constant; $(\Delta E-\$)=9911.250\pm0.063$ MHz and $\alpha^{-1}=137.0356\pm0.0007$.

1. Introduction

The value of the fine structure constant can be obtained using the experimental value of the $2P_{1/2}$–$2P_{3/2}$ energy separation in hydrogen. Several experimentally determined values have been reported [1, 2]. The original value of α, as measured by Lamb and co-workers, disagrees with other determinations (ac Josephson, remeasurement of hydrogen fine structure, etc.) by some 20 ppm (parts per million).

Lamb and co-workers [3] measured the fine structure (fs) of the $n=2$ levels of hydrogen and deuterium using a microwave atomic beam technique. Our experiment is similar to these earlier investigations. An rf electric field is applied to a beam of metastable atomic hydrogen to drive $2S_{1/2}$–$2P_{3/2}$ transitions (see fig. 1). In particular the single hyperfine state β_A $(F=0, m_F=0)$ has been used (see fig. 2). This eliminates to a large degree, corrections for overlapping transitions from the β_B $(F=1, m_F=-1)$ state.

The β_B state has not been used in the present experiment. Robiscoe and Rebane have recently shown that the nonadiabatic transitions (Majorana transitions) [4] used to produce the β_B state from the α states are characterized by an oscillatory velocity distribution which is very sensitive to external magnetic fields. But the β_A state, which is produced by an adiabatic rf transition from the $\alpha^{(-)}$ state $(F=1, m_F=0)$, has a measured approximate v^4 Maxwellian velocity distribution [5] under the present experimental conditions, as shown in figure 3. The velocity distribution results from a combination of the effects due to recoil of the atoms during the electron bombardment production process and collimation of the metastable beam [5, 6, 7]. Since the experimental systematic corrections depends, to some extent, on the metastable velocity distribution, the β_A's, with a measured velocity distribution, were used in preference to the β_B's.

Table 1. *Typical operating conditions*

Source Chamber Pressure	1.2×10^{-4} torr
Bombardment Chamber Pressure	3.2×10^{-6} torr
Detector Chamber Pressure	1.6×10^{-6} torr
Source Pressure	5.8 torr
Oven Temperature	2800 K
Bombardment Voltage	14.0 V
Bombardment Electron Current Density	2.0 mA/cm^2
Electrometer Grid Resistance	5×10^{11} Ω
Estimated Total 2S Yield at Detector	3×10^4 atoms/sec
% β_A Production	37%

The σ-transitions observed were:

β_A-b:

$$2^2S_{1/2}(F=0, m_F=0)-2^2P_{3/2}(F=1, m_F=+1)$$

β_A-d:

$$2^2S_{1/2}(F=0, m_F=0)-2^2P_{3/2}(F=1, m_F=-1).$$

The ranges of the frequencies and the Zeeman magnetic field used for these transitions were 8–11 GHz and approximately 800 G, respectively. Measurements of the rf frequencies and the magnetic fields at the centers of the resonance curves together with

* Supported by the U.S. Atomic Energy Commission.

** Present address: Space Physics Research Laboratories, University of Michigan, Ann Arbor, Mich. 48105.

*** Alfred P. Sloan Foundation Fellow. Present address: Physics Department, Montana State University, Bozeman, Mont. 59715.

the theory of Zeeman splitting were used to determine $(\Delta E - S)$, the energy separation, $2^2S_{1/2}$–$2^2P_{3/2}$.

2. General Procedure and Apparatus

A schematic diagram of the atomic beam machine used in the present experiment is shown in figure 4. A beam of atoms in the 2S metastable state α is produced by thermal dissociation of ground state molecules in a tungsten oven followed by electron bombardment in a 575 G magnetic field. There, due to the motional electric field, the β component is completely quenched. The state selector (flopper)

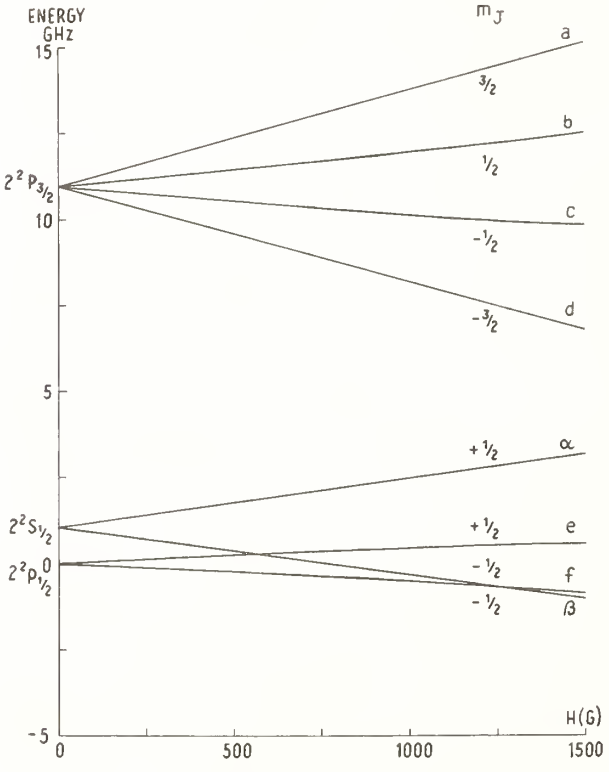


FIGURE 1. *Zeeman diagram of the fine structure of hydrogen* (n = 2).

The separation between $2P_{1/2}$ and $2S_{1/2}$ is the Lamb shift, S. The separation between $2P_{1/2}$ and $2P_{3/2}$ is the fine structure splitting ΔE at zero magnetic field.

is used to regenerate the single hyperfine state β_A from the α state beam by a magnetic dipole transition. In the transition region, an rf cavity located at the center of the Zeeman magnetic field provides a homogeneous localized transverse rf electric field. This field is used to drive a particular 2S–2P transition in the Zeeman magnetic field, resulting in a resonant depletion of the 2S beam. The Zeeman field is parallel to the beam axis to reduce the motional electric fields which can cause asymmetries in the resonance lines and can produce Stark shifts of the 2S and 2P states. The metastable beam is detected with a nickel surface detector. Typical operating conditions for the present experiment are listed in table 1.

The Zeeman magnetic field at the center of the transition region was adjusted to have zero first derivative, and a small positive second derivative. The field was symmetric and constant to within 40 ppm, over a central transition length of 2 cm. Nuclear magnetic resonance probe field readings were within 5 ppm of the average field over the transition length. The field reproducibility was measured to be better than 10 ppm upon repeated insertions of the probe. Field stability was better than 5 ppm over an hour. The frequency stability of the rf oscillator (a Hewlett-Packard 8690A Sweep Oscillator) was measured with an electronic counter (Hewlett-Packard 5245L) and found to be better than 1 ppm for 1-min periods. The electronic counter was accurate to one part in 10^7, as calibrated against WWVB after the final runs.

3. Nature of Data

A total of 84 resonance curve line centers were measured for the $\beta_A - b$ transitions in five runs and 55 for the $\beta_A - d$ transition in five runs. Typical resonance curves are shown in figures 5 and 6 where the Bethe-Lamb theory of the 2S lifetime has been used and an average over a v^4 Maxwellian velocity distribution for the metastable beam has been made.

TABLE 2. *FWHM in Gauss*

$\beta_A - b$ Transitions				$\beta_A - d$ Transitions			
Run No.	F_M (%)	δH_{obs}	δH_{th}	Run No.	F_M (%)	δH_{obs}	δH_{th}
17a	48.3	53.4	52.9	14	43.6	86.3	85.0
18a	43.6	52.2	51.6	17b	46.4	87.8	86.5
19a	46.0	53.8	52.3	18b	35.4	82.8	81.3
20a	49.0	54.6	53.2	19b	43.3	84.3	84.6
21	46.0	53.8	53.1	20b	42.0	85.9	84.2

The full widths at half maximum (FWHM), δH_{obs}, were measured for the resonance curves. In table 2 they are compared with the theoretical FWHM values, calculated using a v^4 velocity distribution. F_M is the maximum fractional amplitude of resonance curve. The discrepancy between the δH_{obs} and δH_{th} is probably due to the approximated velocity distribution of the metastable atoms.

The center of the resonance curve was obtained from the measurements of the beam flop at the two magnetic field points corresponding to $\frac{3}{4}F_M$. The observed line center is located half way between the two field points if the beam flops at these points are the same. The magnetic field values at these points were measured with a proton NMR probe inserted directly in the transition region.

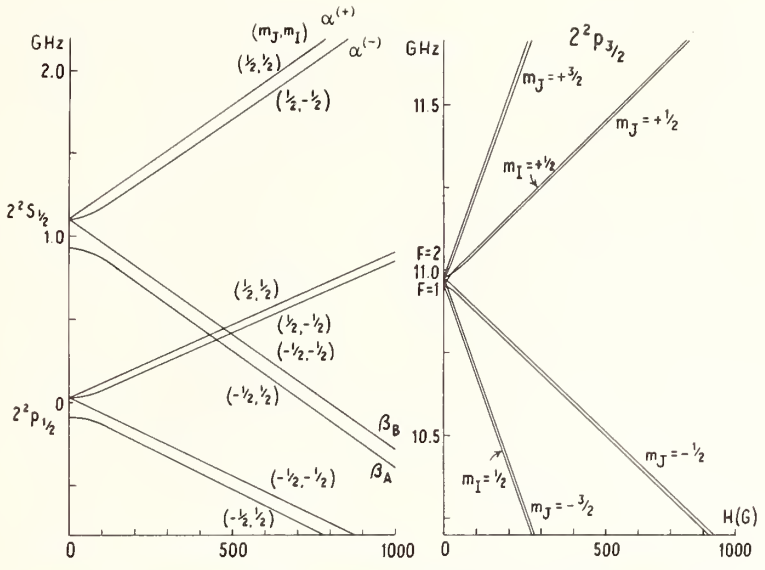


FIGURE 2. *Zeeman diagram of the hyperfine structure of hydrogen* (n=2).
$\Delta\omega$ is the hyperfine energy splitting of the $2S_{1/2}$ state at zero magnetic field.

4. Corrections to the Data

The observed resonance centers are shifted from the true centers by small but significant amounts. These corrections include effects of overlapping transitions, magnetic field variation of the Stark matrix element connecting the states, and the Stark level shift due to the motional and stray fields. Further systematic corrections (Zeeman field inhomogeneity, etc.) are estimated by calculation to be less than 1 ppm of the $(\Delta E - 8)$ value.

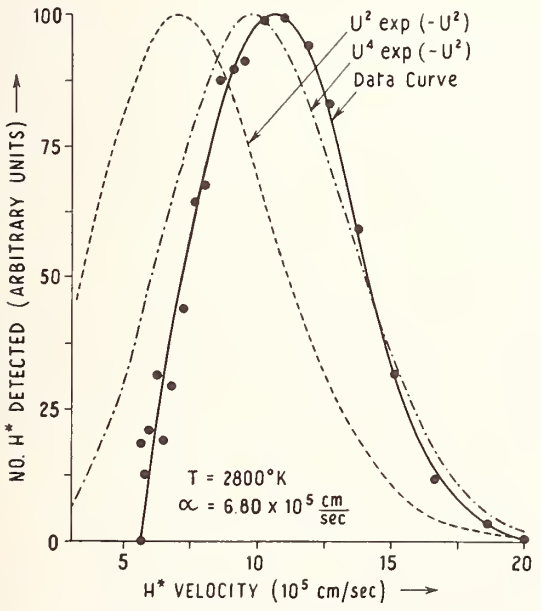


FIGURE 3. *Metastable beam velocity distribution measured by the time of flight technique.*
The dots are experimental points. A Channeltron was used as the metastable atom detector.

4.1. The Overlapping Transitions

The overlapping transitions are resonances that occur at different magnetic fields for a fixed frequency as indicated in figure 7. The first figure shows the resonance curve $F_1(H)$ without such overlapping transitions. The dotted line represents the locus of the midpoints of lines connecting the points which have the same height on the resonance curve. The second figure indicates the distortion that the presence of an overlapping transition $F_2(H)$ introduces into the line shape of $F_1(H)$. For the $\beta_A - b$ transition, the fractional line center shift due to the overlapping transition $\alpha - a$ is typically (88 ± 2) ppm. This raises the observed line center to the high field side. For the $\beta_A - d$ transition, the contribution of the overlapping $\alpha - c$ transition to the fractional line center shift is typically (30 ± 2) ppm. This lowers the observed line center. The fractional β_B impurity in the β_A beam at the working points is $(4.88 \pm 2.56) \times 10^{-3}$. The accidental transition from α to β_B occurs where there is a rapid spatial variation of the magnetic field. This β_B impurity raises the line center (13 ± 7) ppm for the $\beta_A - b$ transition and lowers the line center (60 ± 30) ppm for the $\beta_A - d$ transition. The uncertainty of the β_B impurity introduces an uncertainty in $(\Delta E - 8)$ value. Other overlapping transitions $(\alpha - b, \alpha - d, \beta - e, \beta - f$, etc.) contribute less than 1 ppm to the fractional line center shift; consequently, they are ignored.

4.2. Stark Matrix Element Variation

The $2P_{1/2}$ $(m_j = +\frac{1}{2})$ state is coupled to the $2P_{3/2}$ $(m_j = +\frac{1}{2})$ state by the $\mathbf{L} \cdot \mathbf{S}$ interaction. Conse-

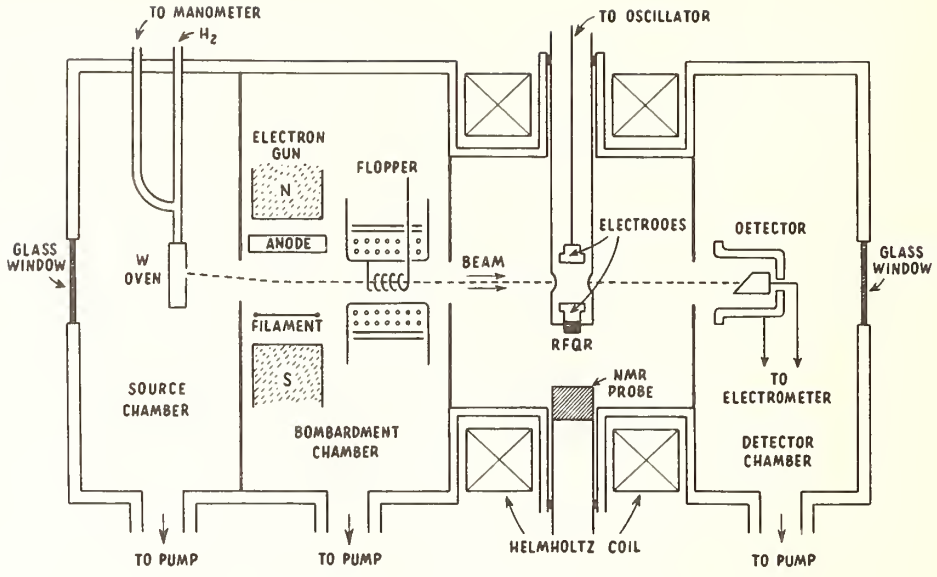


FIGURE 4. *Schematic diagram of the machine.*

quently, $\langle b \mid \mathbf{r} \mid \beta \rangle$ is magnetic field dependent. The matrix element can be approximated in a weak field as

$$| \langle b \mid \mathbf{r} \mid \beta \rangle |^2 = K(H) \mid \langle b \mid \mathbf{r} \mid \beta \rangle |_{H=0}{}^2. \qquad (1)$$

$K(\mathrm{H})$ is 1 for the $\beta_A - d$ transition, while for $\beta_A - b$, it is

$$K(\mathrm{H}) = 3/2(1-\delta_+) \qquad (2)$$

where

$$\delta_+ = (x+\tfrac{1}{3})(1+\tfrac{2}{3}x+x^2)^{-1/2},$$

and

$$x = (g_s - g_L)\mu_0 H/\Delta E.$$

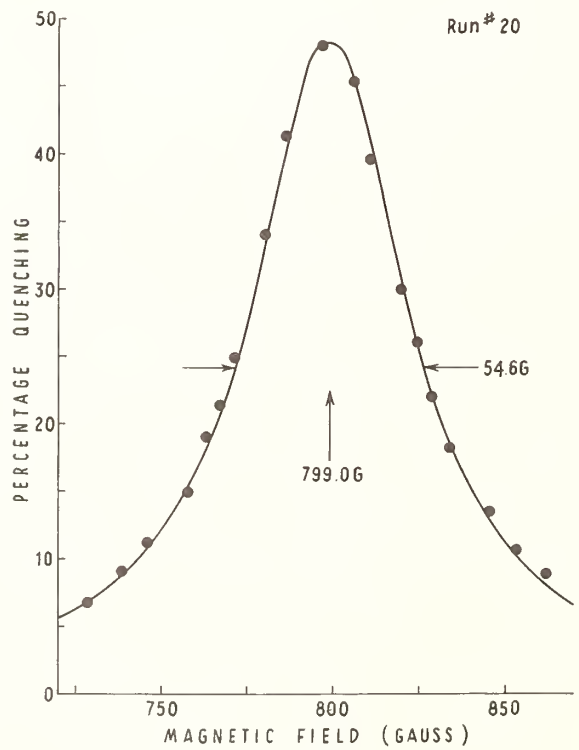


FIGURE 5. *Typical β_A-b panoramic.*
The solid line is the theoretical line shape derived from the Bethe-Lamb theory of the $2S_{1/2}$ lifetime in external fields. The dots are the experimental points.

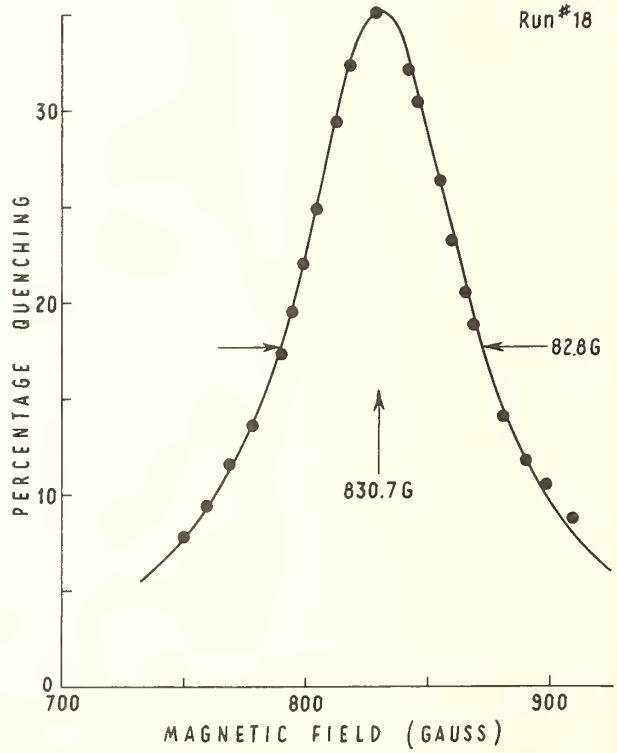


FIGURE 6. *Typical β_A-d panoramic.*
The solid line is the theoretical line shape derived from the Bethe-Lamb theory of the $2S_{1/2}$ lifetime in external fields. The dots are the experimental points.

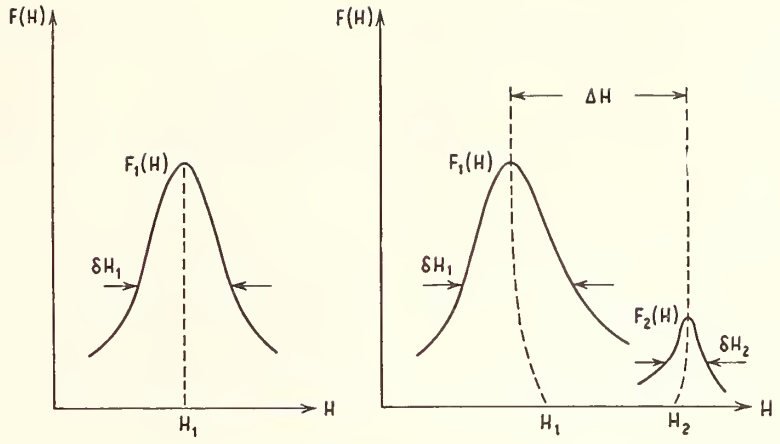


FIGURE 7. *Schematic diagram of an overlapping transition.*

This correction raises the $\beta_A - b$ line center (80 ± 4) ppm.

4.3. Stark Shift

The motional and stray electric fields couple the β state to the 2P states via the quadratic Stark interaction. Since the energy shift of the β state due to the state e is the largest shift and since it is small compared to $(\Delta E - \mathcal{S})$ (less than 1 ppm), the contributions from all other states can be ignored. For the motional field correction, the angle between the beam axis and the magnetic field axis was estimated to be less than 3° due to a machine misalignment. This gives a motional electric field less than 0.4 V/cm. The corresponding energy shift of the β_A state in frequency units is 1.6 kHz. The stray fields were estimated to be 1 V/cm from the beam notch curve [9]. This contribution to the energy shift of the β_A state is 8 kHz. The energy shift due to the motional and stray fields is 10 kHz corresponding to 1 ppm in $(\Delta E - \mathcal{S})$. The directions of the motional and stray fields were not directly determined so these corrections are treated as an uncertainty in the $(\Delta E - \mathcal{S})$ value.

5. Results

To determine $(\Delta E - \mathcal{S})$ in $H(n=2)$, the well established perturbation calculation was adopted [3, 10]. $(\Delta E - \mathcal{S})$ for the $\beta_A - b$ transition and $\beta_A - d$ transition were calculated [8] and are given by, respectively

$$(\Delta E - \mathcal{S})_{\beta_A - b} = f_1 - \tfrac{1}{2} g' \mu_0 H - \tfrac{4}{15} \Delta\omega - \frac{1}{4} \frac{(\Delta\omega)^2}{\mu_0 H} [g_s^{-1} - (75 g_J)^{-1}] - \frac{2}{9} \frac{(g_s - g_L)^2 (\mu_0 H)^2}{\Delta E} + \frac{5}{72} \frac{[(g_s - g_L)\mu_0 H]^3}{(\Delta E)^2} \quad (3)$$

$$(\Delta E - \mathcal{S})_{\beta_A - d} = f_2 + g'' \mu_0 H - \tfrac{1}{5} (\Delta\omega) - \tfrac{1}{4} [(\Delta\omega)^2 / \mu_0 H][g_s^{-1} - (75 g_J)^{-1}] \quad (4)$$

where

$$g' = (g_s + g_L) + \tfrac{1}{3}(g_s - g_L) - \tfrac{1}{36}(g_s - g_L)(\Delta\omega / \Delta E)$$

$$g'' = g_L + \tfrac{1}{72}(g_s - g_L)(\Delta\omega / \Delta E).$$

$\Delta\omega$ = the hyperfine energy splitting in the 2S state of hydrogen.

The numerical values used for the physical constants in the above equations are [11, 12, 13]:

$$g_s = 2.00232$$

$$g_L = [1 - (m_e/m_p)] = 0.999455$$

$$g_J = 1.33374 \text{ for } J = 3/2$$

$$\mu_0 H = g_s^{-1}(g_s(\text{free})/g_p(\text{water}))\omega = 328.732465\ \omega.$$

where ω is the proton resonance frequency in MHz corresponding to the magnetic field H.

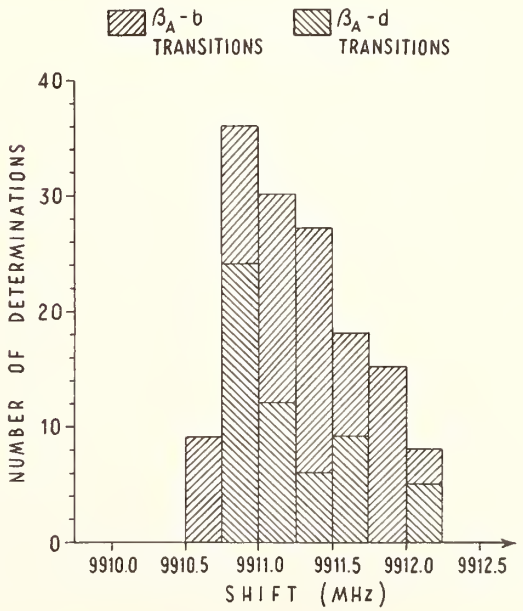


FIGURE 8. *Histogram of the* $(\Delta E - \mathcal{S})$ *distribution from 139 line center measurements.*

Table 3. *Summary of experimental uncertainties (in MHz)*

Statistical Average Deviation	0.052
Background Correction	0.020
β_B Impurity	0.025
Stark Shift	0.010
Vector Sum	0.063

The results from the present experiments are

$$(\Delta E - \mathcal{S}) = 9911.255 \pm 0.059 \text{ MHz for the } \beta_A - b \text{ transitions;}$$

$$(\Delta E - \mathcal{S}) = 9911.242 \pm 0.090 \text{ MHz for the } \beta_A - d \text{ transitions} \quad (6)$$

including statistical errors, overlap corrections and Stark matrix element variation. The distribution of the $(\Delta E - \mathcal{S})$ values obtained from 139 line center measurements are shown in the histogram in figure 8. Since the two distributions are not clustered about different points, the individual data were treated with equal weights. The weighted average from all data is

$$(\Delta E - \mathcal{S}) = 9911.250 \pm 0.052 \text{ MHz} \quad (7)$$

where the quoted precision is one average deviation from the mean.

Including all sources of experimental uncertainty, shown in table 3, the final value of $(\Delta E - \mathcal{S})$ from the present experiments is

$$(\Delta E - \mathcal{S}) = 9911.250 \pm 0.063 \text{ MHz.} \quad (8)$$

The quoted experimental uncertainty for the final result is a 67 percent confidence level. The individual uncertainties were summed quadratically. The uncertainty for the background correction came from the process of normalizing the resonance curves as indicated in ref [3].

Four experimental $(\Delta E - \mathcal{S})$ values in $H(n=2)$ have been reported and are listed in table 4. The experiments done by Robiscoe, Williams and Cosens [14] were similar to the present experiments. The primary difference was that they used atoms in the β_B state, rather than β_A. The Kaufman et al. value [15] of $(\Delta E - \mathcal{S})$ is 0.13 MHz higher than ours and does not agree with our value even within the experimental uncertainty. Cosens and Vorburger [16] used an atomic beam radio frequency method similar to that used in the present experiment, near a Zeeman field of 400 G. Their value is 0.08 MHz lower than ours, but agrees within experimental uncertainty.

In order to determine the fine structure interval, ΔE, the revised value of the $2S_{1/2}$–$2P_{1/2}$ energy interval [5], $\mathcal{S}_H = (1057.90 \pm 0.10)$ MHz was used. Recently, Appelquist and Brodsky [17] found a calculational correction for the Lamb shift. Their corrected value, $\mathcal{S}_H(\text{theory}) = (1057.91 \pm 0.16)$ MHz, agrees almost exactly with the revised value of

Table 4. $(\Delta E - \mathcal{S})$ *in* $H(n=2)$

9911.0±0.4 MHz	Robiscoe, Williams and Cosens [14]
9911.38±0.03 MHz	Kaufman et al. [15]
9911.175±0.042 MHz	Cosens and Vorburger [16]
9911.250±0.063 MHz	Present Experiment

$\mathcal{S}_H(\text{exp.})$ which is proposed in ref. [5] (including the kinematic corrections to the atomic beam distribution). The fine structure interval, ΔE, obtained is 10969.15±0.12 MHz, and the corresponding value for α^{-1} is 137.0356±0.0007. The quoted uncertainty of ΔE is the vector sum of one average deviation in $(\Delta E - \mathcal{S})$ and two average deviations in S_H. Our value of α^{-1} is about 20 ppm less than the value derived from the fine structure splitting in deuterium as measured by Triebwasser, Dayhoff, and Lamb [3], $(\alpha^{-1})_{TDL} = 137.0388 \pm 0.0006$. Our value of α^{-1} agrees with the recent α^{-1} value from measurements of $2e/h$ using the Josephson effect [2], $(\alpha^{-1})_{2e/h} = 137.03608 \pm 0.00026$, and with the optical level crossing value [18], $(\alpha^{-1})_{OLC} = 137.0354 \pm 0.0007$.

6. References

[1] Robiscoe, R. T., Cargese Lectures in Physics, edited by M. Levy (Gordon and Breach, Inc., New York, 1968), pp. 3–53.

[2] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[3] This series of six papers on the fine structure of hydrogen is as follows: (H–I): Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **79**, 549 (1950). (H–II): Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **81**, 222 (1951). (H–III): Lamb, W. E., Jr., Phys. Rev. **85**, 259 (1952). (H–IV): Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **86**, 1014 (1953). (H–V): Triebwasser, S., Dayhoff, E. S., and Lamb, W. E., Jr., Phys. Rev. **89**, 98 (1953). (H–VI): Dayhoff, E. S., Triebwasser, S., and Lamb, W. E., Jr., Phys. Rev. **89**, 106 (1953).

[4] Robiscoe, R. T., and Rebane, T., private communication.

[5] Robiscoe, R. T., and Shyn, T. W., Phys. Rev. Letters **24**, 559 (1970).

[6] Pearl, J. C., Donnelly, D. P., and Zorn, J. C., Phys. Letters **30A**, 145 (1969).

[7] Pearl, J. C., Ph.D. thesis, The University of Michigan (1970).

[8] Shyn, T. W., Ph.D. thesis, The University of Michigan (1969).

[9] The beam notch is the sharp decrease of the β beam near 575 G (the β-e crossing point) arising from quenching by motional and stray electric fields. See Ref. 8 and R. T. Robiscoe, Phys. Rev. **168**, 4 (1968).

[10] Robiscoe, R. T., Phys. Rev. **138**, A22 (1965).

[11] Wilkinson, D. T., and Crane, H. R., Phys. Rev. **130**, 852 (1967).

[12] Myint, T., Kleppner, D., Ramsey, N. F., and Robinson, H. G., Phys. Rev. Letters **17**, 405 (1966).

[13] Rich, A., Phys. Rev. Letters **20**, 967 (1968).

[14] See DuMond, J. W. M., Z. Naturforsch. **219**, 70 (1966).

[15] Kaufman, S. L., Lamb, W. E., Jr., Lea, K. R., and Leventhal, M., Phys. Rev. Letters **22**, 507 (1969).

[16] Cosens, B. L., and Vorburger, T. V., Phys. Rev. Letters **23**, 1273 (1969). Also to be published in Phys. Rev.

[17] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970).

[18] Metcalf, H., Brandenberger, J. R., and Baird, J. C., Phys. Rev. Letters **21**, 165 (1968). See also these Proceedings.

A Measurement of ΔE-$\mathcal{S}$ in H and a Correction to the Measured Value of $\mathcal{S}$ in D*

T. V. Vorburger and B. L. Cosens*****

Yale University, New Haven, Conn. 06520

The $2^2S_{1/2}$–$2^2P_{3/2}$ separation $\Delta E-\mathcal{S}$ in hydrogen has been measured by a microwave atomic beam technique. The result obtained from four independent transitions is $\Delta E-\mathcal{S}=9911.173\pm0.042$ MHz. We combine this result with the recent determination of the Lamb shift $\mathcal{S}$ by Robiscoe to obtain a value for the fine structure splitting ΔE and a value for the fine structure constant $\alpha^{-1}=137.0358(5)$. We also report a correction in a previous measurement of the Lamb shift $\mathcal{S}$ in D resulting from a beam velocity distribution measurement recently made by Robiscoe and Shyn. The corrected value is $1059.28\pm.06$ MHz.

Key words: Deuterium fine structure; fine structure constant; hydrogen fine structure.

1. Introduction

The present experiment is the last in a series of experiments that has remeasured the fine structure of hydrogen for $n=2$ (fig. 1). The present experiment is a remeasurement of $\Delta E-\mathcal{S}$, the $2^2P_{3/2}$–$2^2S_{1/2}$ energy splitting in H [1]. Previously, Robiscoe and Cosens measured the Lamb shift, the $2^2S_{1/2}$–$2^2P_{1/2}$ splitting in H [2], and Cosens measured the Lamb shift in D [3]. These experiments and a great number of others including the original work of Lamb and co-workers [4] have provided experimental tests of the theory of quantum electrodynamics. The present measurement of $\Delta E-\mathcal{S}$ may be combined with the Robiscoe value for $\mathcal{S}$ in hydrogen to yield a value for the fine structure splitting ΔE and a value for the fine structure constant α.

2. The Experiment

The $2^2S_{1/2}$ states labeled α and β are metastable. They live for about $\frac{1}{8}$ s before decaying to the ground state by two photon emission. All of the 2P states (a, b, c, d, e, f) have a short lifetime of about 1.6×10^{-9} s. The experiment is done by producing a beam of hydrogen atoms in a pure hyperfine sublevel of the β state, i.e., in either the $m_I=+\frac{1}{2}$ sublevel or the $m_I=-\frac{1}{2}$ sublevel (fig. 2). This beam enters a homogeneous magnetic field of about 400 G where it is subjected to an rf electric field and undergoes electric dipole transitions to the short lived b or d levels in the $2^2P_{3/2}$ state. The transitions are observed as a decrease in the metastable beam signal. Precise measurements of the rf frequency and the magnetic field together with an extrapolation to zero magnetic field along the Zeeman levels yields a value for $\Delta E-\mathcal{S}$. We have measured four transitions as shown in table 1: βb^-, βb^+, βd^-, βd^+. These yield four independent values for $\Delta E-\mathcal{S}$.

The previous experiments [2, 3] measured the Lamb shift in H and D by indueing dc transitions between the levels β and e (fig. 1). This transition has zero frequeney at the β–e crossing point of 575 G. A precise measure of the crossing point magnetic field yielded the value of $\mathcal{S}$. This erossing point phenomenon is relied on heavily in the present experiment also.

3. The Apparatus

Nearly all of the techniques in the experiment have been described earlier by Lamb et al., and by Robiscoe and Cosens.

Figure 3 shows the apparatus. A beam of hydrogen atoms is produced in a tungsten oven and is excited to the 2S state by electron bombardment. The beam proceeds toward the detector. The β state atoms are quenched as they are produced in the excitation region by a motional eleetric field. The beam enters the state selector or flopper, which can be operated in one of two modes to populate the β^+ or the β^- state. The device is a solenoid which supplies a low magnetic field (0–15 G). At the center of the solenoid is a small rf coil. The β^+ state (see fig. 2) is populated by adjusting the magnetic field to a few gauss and inducing magnetic dipole transitions from the α^- state to the β^+ state with an rf frequency of about 179 MHz. The β^- state atoms are produced by Majorana transitions. The magnetic field of the flopper is set equal to zero where the α^+, α^-, and β^- levels are all degenerate. The beam atoms are re-

* Work supported by the National Science Foundation.

** Present address: Physics Department, University of Delaware, Newark, Dela. 19711.

*** Present address: Physics Department, St. Olaf College, Northfield, Minn. 55057.

TABLE 1. *Results of the four transitions*

Transition	Approx. frequency (GHz)	Sum of corrections (MHz)	Systematic error (MHz)	Random error (MHz)	$\Delta E-\delta$ (MHz)
βb^-	10.8	−0.057	±0.055	±0.074	9911.281±0.092
βb^+	10.8	−0.070	±0.039	±0.076	9911.144±0.085
βd^-	9.4	+0.016	±0.022	±0.073	9911.196±0.076
βd^+	9.4	+0.009	±0.053	±0.065	9911.084±0.084

Weighted Mean: $\Delta E-\delta=9911.173\pm0.042$ MHz

$\Delta E=(\Delta E-\delta)+\delta=(9911.173\pm0.042)+(1057.896\pm0.063)$

$=10\ 969.069\pm0.076$ MHz

$\alpha^{-1}=137.0358\ (5)$

oriented when they pass through this region of zero field so that the beam emerges with some of the atoms populating the β^- state.

The beam enters the Helmholtz coil, which supplies a homogeneous magnetic field of about 400 G parallel to the beam axis. At the center is an rf cavity where the transitions take place. The direction of the rf electric field is perpendicular to the static magnetic field.

The beam now passes through the quencher, which serves as a switch for the β state metastable atoms. It is yet another solenoid, which supplies an axial magnetic field of 575 G, the field of the β–e crossing point. At the center a pair of electrodes supplies an electric field of several volts per cm. When the quencher is turned on, essentially all of the β state atoms are quenched by dc transitions to the e state.

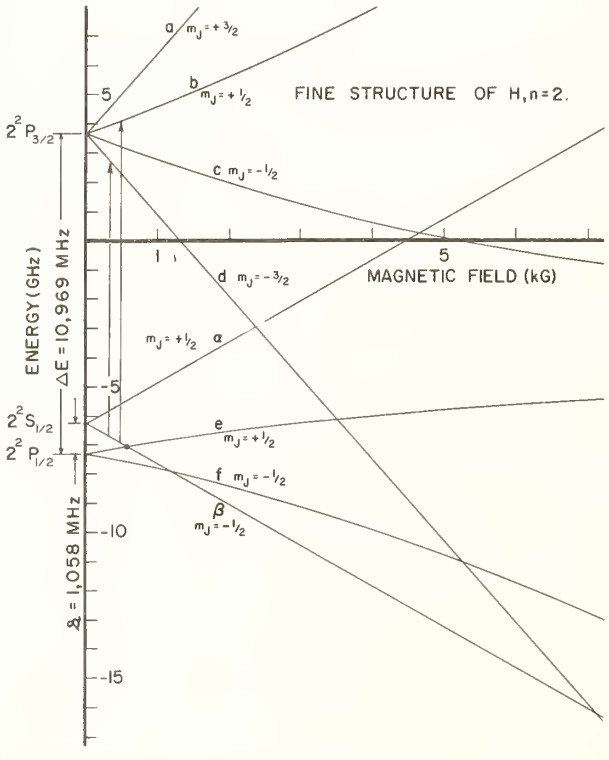


FIGURE 1. *Zeeman diagram for* H, n = 2.
ΔE and δ are shown. The transitions βb and βd are represented by arrows. The β-e crossing is also shown.

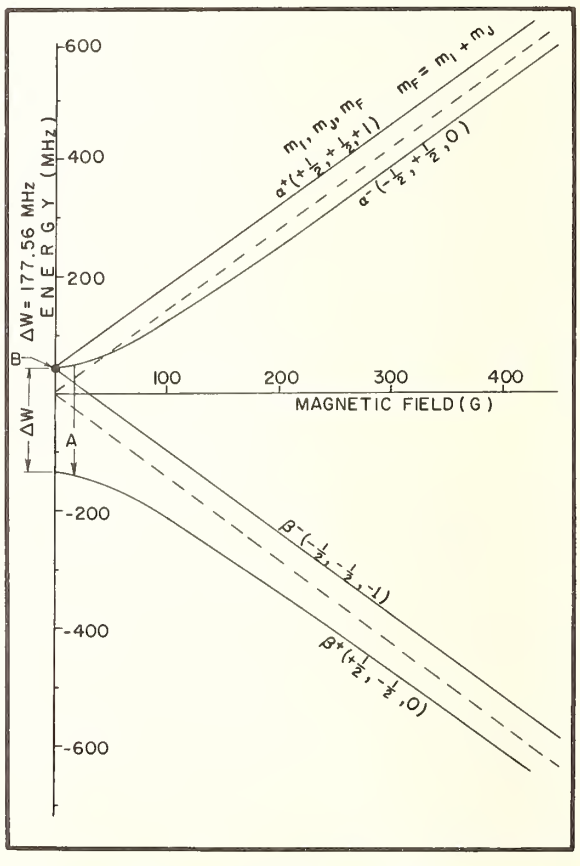


FIGURE 2. *Hyperfine structure of the* $2^2S_{1/2}$ *state.*
Arrow A represents the rf transition used to populate the β^+ level. Point B represents the Majorana transition used to populate the β^- level.

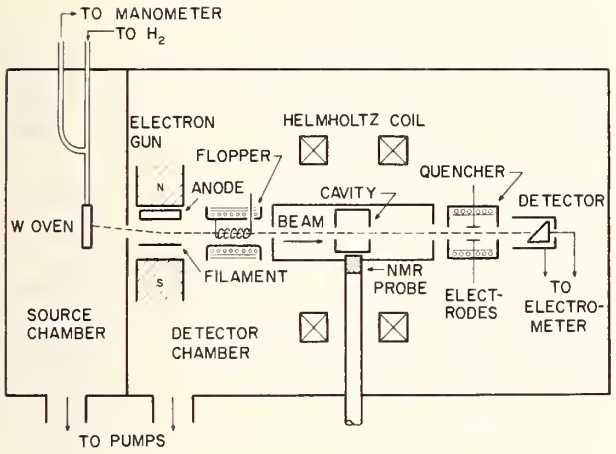


FIGURE 3. *Schematic diagram of the apparatus.*

4. Use of the Quencher

Figure 4 shows how the data is taken and why we use the quencher. The top graph shows a sample rf quenching signal. The rf frequency is held constant, and the magnetic field is swept through the resonance. We see a βb^- resonance, whose center we want to determine, and an overlapping αa resonance with a central field of about 670 G. When the quencher is turned on, we get the second graph. There are very few β state atoms left in the beam, so we see only the αa resonance. Subtracting the two curves yields an almost pure βb^- rf resonance. The central magnetic field of this resonance line is then measured with a precision of ~100 ppm; while the rf frequency is held constant to ~1 ppm.

There is a similar overlap of the βd resonance by the αc resonance. The quencher is therefore used in the measurements on all of the transitions.

5. Theoretical Line Shape and Corrections

We now need to make a correction for the little remnant of α resonance which we saw in figure 4. A plot of the data corrected for this asymmetry yields the final experimental line shape shown in figure 5. The curve is a theoretical line shape normalized to the experimental maximum. It has been derived from the Bethe-Lamb theory for the lifetime of the 2S state perturbed by an rf electric field. The line shape must also be averaged over the velocity distribution of the beam. It was originally thought that the distribution had a v^2 Maxwellian dependence [4]. However, a recent measurement by Robiscoe and Shyn [5] has shown that the actual distribution is closer to a v^4 Maxwellian. We have assumed a v^4 distribution in these calculations.

There are many other effects (about twelve in all), which asymmetrize the resonance line and shift its center. Some of these were first discussed by Lamb [4] such as the variation of the rf matrix element with magnetic field. We have to make a 0.1 MHz

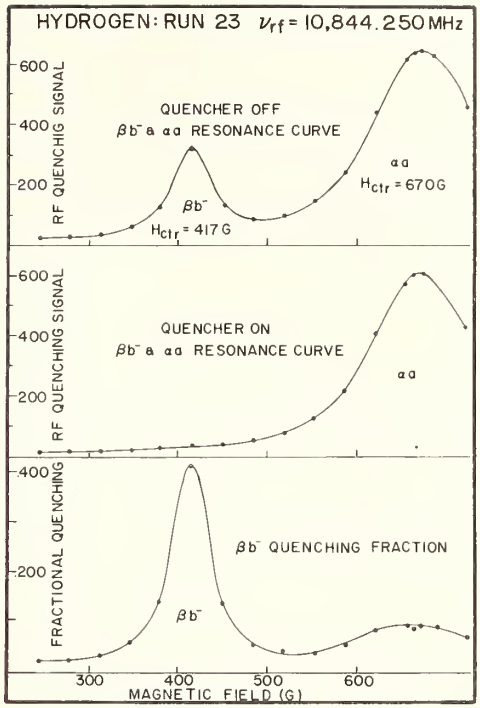


FIGURE 4. *RF resonances for Run 23.*
(a) With quencher off: Graph shows βb^- and overlapping αa resonances.
(b) With quencher on: The β state atoms are quenched. Graph shows only αa resonance.
(c) RF quenching fraction given by the difference between (a) and (b) divided by the total β signal.

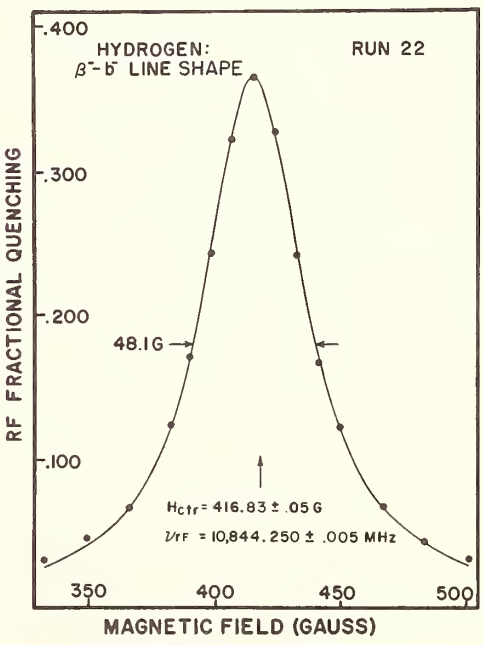


FIGURE 5. *βb^- quenching fraction for Run 22.*
The experimental points are corrected for α overlap. The theoretical curve is derived using the Bethe-Lamb theory for the lifetime of the 2S state under a perturbing electric field. The line shape is averaged over a v^4 velocity distribution.

correction due to this effect on the βb transitions. Another effect, the distortion of the beam velocity distribution by motional and stray electric fields, was discovered by Robiscoe [6]. For the βd^+ transition this asymmetry causes $\sim$0.04 MHz change in $\Delta E - \delta$. Many of these corrections are extremely small (<0.01 MHz) such as the rf Stark shift and the βd overlap on the βb transition. Table 1 (col. 3) shows average values for the sum of the corrections for each transition.

6. Results

After all the corrections are applied, the transition frequencies are extrapolated to zero magnetic field by using a computer diagonalization of Brodsky and Parsons' Hamiltonian for H, $n=2$ [7]. Brodsky and Parsons state that their analysis of the Zeeman structure is intrinsically accurate to 0.001 MHz.

The results for $\Delta E - \delta$ are shown in table 1. The fourth column shows the estimated 70 percent uncertainty in the corrections, and the fifth column shows the statistical standard deviation. The final results are 9911.281 MHz, 9911.144 MHz, 9911.196 MHz, and 9911.084 MHz. The total error is the quadratic sum of the statistical and the systematic error. The final value for $\Delta E - \delta$ is a weighted mean average of the four: 9911.173$\pm$0.042 MHz. The quoted error is one standard deviation of the mean. The external standard deviation of the four transition values is $\pm$0.040 MHz.

We calculate ΔE by adding the Robiscoe value for the Lamb shift of 1057.896$\pm$0.063 MHz [5]. The value of ΔE is then equal to 10 969.069$\pm$0.076 MHz. This gives a result for $\alpha^{-1}=137.0358(5)$.

7. Discussion

The results for $\Delta E - \delta$ may be compared with the result of Shyn, Williams, Robiscoe, and Rebane [8], whose apparatus is similar to the present one. They have measured βb^+ and βb^- transitions at a magnetic field of about 850 G. Their final average from 115 line centers is 9911.250$\pm$0.063 MHz. The difference between the two results is 0.077$\pm$0.076 MHz. Another result comes from the group of Kaufman, Lamb, Lea, and Leventhal [9], who have measured αa and αb transitions at a field of about 1600 G and obtained a result of 9911.38$\pm$0.03 MHz. The disagreement between our result and theirs is about 4 standard deviations and is not understood.

The result for ΔE may be compared with that of Metcalf, Brandenberger, and Baird [10], who make a determination of ΔE by measuring the e–d level crossing in the 2P state. Their result of 10 969.13$\pm$ 0.12 MHz differs by 0.06$\pm$0.14 MHz.

The present result for α may be compared with several other experiments. Figure 6 shows some of these including the important measurement of e/h by Parker, Taylor, Finnegan, Denenstein, and Langenberg [11] and the results from the measurements of the hydrogen ground state hyperfine splitting [12]. Also shown are the results from the muonium hyperfine structure measurements [13] and from H and D fine structure.

Since the latest calculation of the Lamb shift by Appelquist and Brodsky [14], the agreement between theory and experiment in H fs has improved greatly. There should now be more confidence in the H fs experiments. However, it seems unlikely that these experiments can really help towards deter-

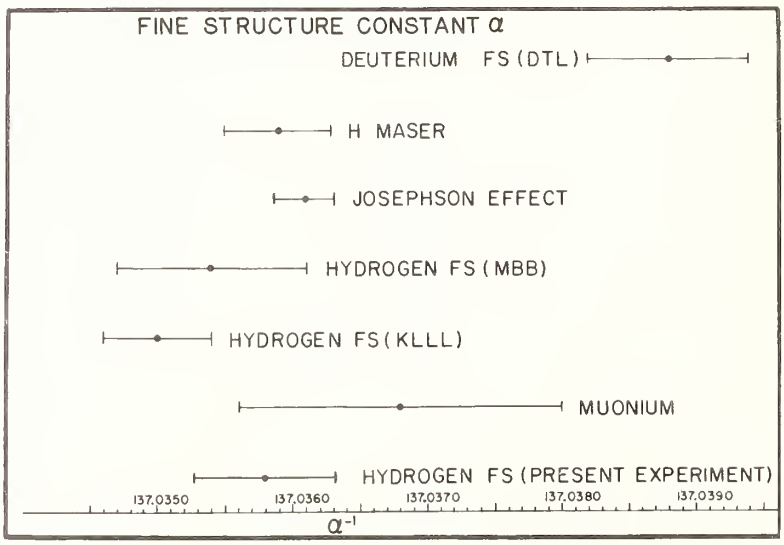


FIGURE 6. *The fine structure constant determined from seven experiments including the present one.*

mining a value for α since it appears difficult to reduce the experimental error much further. We feel that these measurements can serve at best as confirming the more accurate methods of measuring the fine structure constant.

8. Correction to the Lamb Shift in Deuterium

As a final note we would like to mention a correction to the measurement by Cosens of the Lamb shift in deuterium [3]. The correction results from the measurement of a v^4 velocity distribution by Robiscoe and Shyn. The new result is 1059.28 ± 0.06 MHz, which puts it in slightly better agreement with the theoretical calculation.

9. Acknowledgment

We wish to express our thanks to Professor William Lichten for his support during this series of experiments and for his many valuable contributions throughout the course of this work.

10. References

[1] Cosens, B. L., and Vorburger, T. V., Phys. Rev. Letters **23**, 1273 (1969); Cosens, B. L., and Vorburger, T. V., Phys. Rev. **A2**, 16 (1970).
[2] Robiscoe, R. T., Phys. Rev. **138**, A22 (1965); Robiscoe, R. T., and Cosens, B. L., Phys. Rev. Letters **17**, 69 (1966).
[3] Cosens, B. L., Phys. Rev. **173**, 49 (1968).
[4] Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **79**, 549 (1950), and **81**, 222 (1951); Lamb, W. E., Jr., Phys. Rev. **85**, 259 (1952); Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **86**, 1014 (1952); Triebwasser, S., Dayhoff, E. S., and Lamb, W. E., Jr., Phys. Rev. **89**, 98 (1953); Dayhoff, E. S., Triebwasser, S., and Lamb, W. E., Jr., Phys. Rev. **89**, 106 (1953).
[5] Robiscoe, R. T., and Shyn, T. W., Phys. Rev. Letters **24**, 559 (1970).
[6] Robiscoe, R. T., Phys. Rev. **168**, 4 (1968).
[7] Brodsky, S. J., and Parsons, R. G., Phys. Rev. **163**, 134 (1967).
[8] Shyn, T. W., Williams, W. L., Robiscoe, R. T., and Rebane, T., Phys. Rev. Letters **22**, 1273 (1969).
[9] Kaufman, S. L., Lamb, W. E., Jr., Lea, K. R., and Levanthal, M., Phys. Rev. Letters **22**, 507 (1969).
[10] Metcalf, H., Brandenberger, J. R., and Baird, J. C., Phys. Rev. Letters **21**, 165 (1968); J. R. Brandenberger, thesis, Brown University, 1968 (unpublished).
[11] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969); Finnegan, T. F., Denenstein, A., and Langenberg, D. N., Phys. Rev. Letters **24**, 738 (1970).
[12] Crampton, S. B., Kleppner, D., and Ramsey, N. F., Phys. Rev. Letters **11**, 338 (1963); Vessot, R., et al., IEEE Trans. on Inst. and Meas. **IM-15**, 165 (1966).
[13] Thompson, P. A., Amato, J. J., Crane, P., Hughes, V. W., Mobley, R. M., zu Putlitz, G., and Rothberg, J. E., Phys. Rev. Letters **22**, 163 (1969).
[14] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970).

DISCUSSION

R. T. Robiscoe: What sort of agreement did you get between theory and experiment for the linewidths in this experiment? And how accurately do you think the $2P$ lifetime could be measured by looking at the line widths?

T. V. Vorburger: The agreement between theory and experiment is very good. The ratio of experimental to theoretical line width is something like 0.992 for the β-b transitions and 0.999 for the β-d transitions. We also did a measurement of the height of resonance the curve at the working points, and theory versus experiment there was 1.005 and 0.995 for β-d and β-b respectively.

R. T. Robiscoe: You're speaking of ratios?

T. V. Voburger: Ratios. Right. Of one to the other. And the statistical spread is very good. However, there is a problem in the normalization which we have from the α overlap—There are two ways of measuring the height of the overlapping α resonance, the one way you saw here in figure 4 and another way which disagrees; it gives a lot smaller result for the resonance height. Because of that uncertainty, there is a consequent uncertainty in the normalization of our beta resonance, and so I would say there's probably plus or minus a percent and a half there. As far as determining the lifetime of the $2P$ states, I haven't given it any thought, but it's probably a couple of percent. And that's not so good I think. It's not a good experiment.

H. Metcalf: I have two questions. First, you have listed a table of corrections to each of the numbers, and the sum of those 12 numbers was of the order of 0.05 megacycles. I guess I should say megahertz here. What are the size of the numbers that went into that sum? Were some of them, many of them, larger than that?

T. V. Vorburger: The largest one as I remember is the one I mentioned, the rf matrix element variation, which is 0.1 megahertz.

H. Metcalf: All the others are smaller than that?

T. V. Vorburger: Yes. The velocity distribution one is about 0.04 MHz. The α-a overlap which hurts the β-b resonance is something like 0.05 MHz.

H. Metcalf: Second question. I gather from what you have said that after you have applied all these corrections to your data you then fit this data with a theoretical curve, and I presume that you fit it with a symmetric curve. Have you ever tried to fit it with an asymmetric curve?

T. V. Vorburger: No. Oh, no, we don't fit it. We have never fitted a curve.

H. Metcalf: You work only on the slopes?

T. V. Vorburger: Yes. Right.

H. Metcalf: Okay.

T. V. Vorburger: However, the theoretical curve that we use for comparison only is a symmetric one.

H. Metcalf: You don't fit it with a curve if you work on the slopes?

T. V. Vorburger: Right.

"It appears that all is not well with quantum electrodynamics." . . . Taylor, Parker, and Langenberg, Rev. Mod. Phys., 1969.

In order to test QED,
We measured 2S to 2P
But now TPL
Just say "All is not well
With Kaufman, Lamb, Leventhal, and Lea."

K. R. Lea

Measurement of the $2^2S_{1/2}$-$2^2P_{3/2}$ Interval in Atomic Hydrogen*

S. L. Kaufman, W. E. Lamb, Jr., K. R. Lea, and M. Leventhal***

Yale University, New Haven, Conn. 06520

The interval $2^2S_{1/2}$–$2^2P_{3/2}$ in hydrogen, denoted $(\Delta E-S)_H$, has been determined through measurements on the two transitions αa ($M_J=+\frac{1}{2}\rightarrow+\frac{3}{2}$) and αb ($M_J=+\frac{1}{2}\rightarrow+\frac{1}{2}$). Details of the microwave-optical experiment are described. The line shape theory is briefly discussed, and the data is summarized. An outline of the treatment of several small corrections is included. The weighted average gives $(\Delta E-S)_H=9911.377\pm0.026$ MHz. This is compared with a recent theoretical result.

Key words: Fine structure constant; fine structure of hydrogen; Lamb shift.

1. Introduction

Studies of atomic fine structure, particularly of one-electron atoms, have been undertaken in the past 25 years with the aim of testing the Dirac theory of the hydrogen atom, and of checking the radiative corrections to the Dirac theory provided by quantum electrodynamics [1]. Under the mutual stimulus of theoretical and experimental developments, improved precision has been achieved over the years in both measured and calculated values of hydrogenic fine structure. We report here the results of one experiment which measured the interval $2^2P_{3/2}$–$2^2S_{1/2}$ in hydrogen, denoted $\Delta E-S$, to a precision of 3 parts per million. The experimental method we employed was originally selected so as to differ substantially from the atomic beam methods previously used to measure this interval. The motivation for this decision was that of providing a completely independent measurement, requiring entirely new considerations of systematic errors. Whilst the experimental hardware has evolved substantially from that originally conceived, it remains true that the method differs in several important respects from those used by other workers in the field. To distinguish it from these atomic beam methods, ours has been called "the bottle method."

* Research supported in part by the U.S. Air Force Office of Scientific Research.
** Present address: I Physikalisches Institut der Univ. Heidelberg, Heidelberg, Germany.
*** Present address: Bell Telephone Laboratories, Murray Hill, New Jersey 07974.

2. The Experiment

Three steps which summarize the method are shown in figure 1:

(i) excitation of the metastable $2S_{1/2}$ state of atomic hydrogen by electron bombardment of molecular hydrogen;
(ii) stimulation of electric dipole transitions between $2^2S_{1/2}$ and $2^2P_{3/2}$ by an applied, resonant radio-frequency field;
(iii) detection of the ensuing Lyman-α radiation emitted by atoms in the short-lived $2^2P_{3/2}$ state upon spontaneous decay to the ground state.

This sequence of events is accomplished in the apparatus shown in figure 2. An electron beam traverses a section of X-band waveguide, to which low pressure hydrogen is admitted. The metastable hydrogen atoms created there are subject to the rf field inside the waveguide, and those undergoing an rf transition to the $^2P_{3/2}$ state subsequently emit Lyman-α radiation. This apparatus is located in a uniform magnetic field, which serves several purposes, one being to confine the electron beam to a well defined column as it crosses from cathode to anode. The photodetector, which is well removed from the magnet in addition to being magnetically shielded, is optically coupled to the radiating atoms by means of an internally polished metal ellipsoidal light pipe. The foci are arranged to be one at the

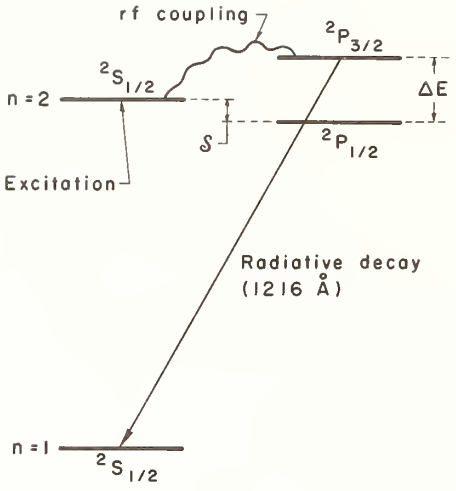


FIGURE 1. *Principle of the experiment.*

center of the waveguide interaction region, and the other at the window of the detector. The detector is an ionization chamber, with a magnesium fluoride window, containing low pressure nitric oxide and carbon dioxide. It responds in the wavelength range 1180 to 1343 Å with an overall quantum efficiency reportedly as large as 50 percent at the Lyman-α wavelength, 1216 Å. The photocurrent is amplified and fed to a lock-in detector, where the signal is separated out of the 100 times larger total light reading by the well-established method of synchronous detection. For this purpose, the microwave field is square-wave modulated on and off about 40 times a second. It follows that the *signal*, by which we mean the increase in intensity of 1216 Å radiation caused by the microwave transitions, is similarly

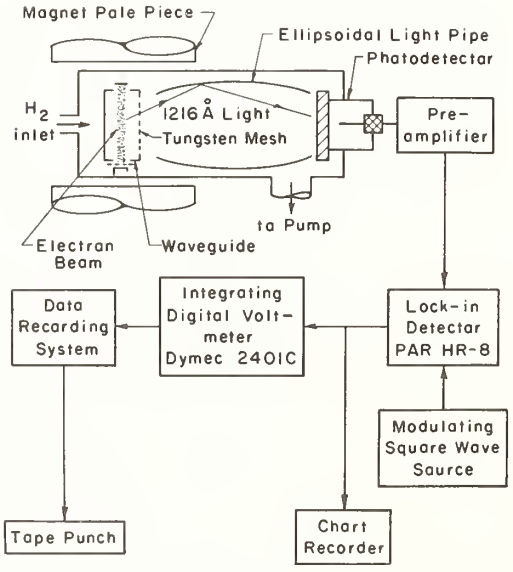


FIGURE 2. *Schematic diagram of the apparatus.*

modulated. The rectified signal, which appears at the output of the lock-in detector, is supplied to the integrating digital voltmeter. The voltmeter integrates the signal over a 10 s interval, and its reading is then recorded on punched paper tape.

Other items of equipment, not all explicitly indicated in figure 2, include the following:

(i) a magnetic field regulator, based on the Hall effect, which reduces field drift to about 20 ppm in 30 min, an entirely acceptable level;
(ii) a proton resonance system for the precise measurement of magnetic field at the location of an NMR probe in the magnet gap;
(iii) a vacuum pumping station, incorporating a Welch turbo-molecular pump, operating continuously, together with a capillary tube device for controlling the hydrogen inlet to the system, and Pirani and ionization gauges for pressure measurement;
(iv) an electron gun, comprising a tungsten filament and a grid adjacent to an aperture in one wall of the waveguide section, together with an anode plate just beyond an aperture in the opposite wall; the current reaching the anode is electronically regulated, and is stable to better than 0.02 percent per half hour;
(v) a microwave system which includes an oscillator (synchronized to a harmonic of a crystal reference) followed by a PIN diode modulator, which delivers square wave modulated microwave power to the experimental region; the incident power is levelled by a feedback loop in which changes in the power level sampled by a crystal detector are automatically compensated for by adjustments to the bias current of the modulator.

With the foregoing equipment, a large body of experimental data was accumulated on three microwave resonanees. Figure 3 shows a plot of transition frequency versus magnetic field for the ten allowed transitions between the $n=2$ levels of hydrogen. The notation is that of Lamb and Retherford [2], wherein the Zeeman sublevels of the $^2S_{1/2}$, $^2P_{3/2}$ and $^2P_{1/2}$ states are labelled α, β, a through d, and e, f respectively. Note that hyperfine structure is not included in this figure. The dotted line shows the position of the electron cyclotron resonance, and the dashed lines show allowed microwave transitions in the $n=3$ levels of hydrogen, to all of which our equipment is sensitive. The natural width of the transitions is 100 MHz (dependent on the inverse lifetimes of the participating states), but in practice this is increased by an rf-dependent broadening, not to mention the hyperfine splitting. To avoid the complications of overlaps from other transitions which are noticeably more numerous in the lower frequency area in figure 3, and for other reasons, it was decided to make measurements on the relatively well separated transitions in the upper portion of figure 3. These six transitions are all straight-

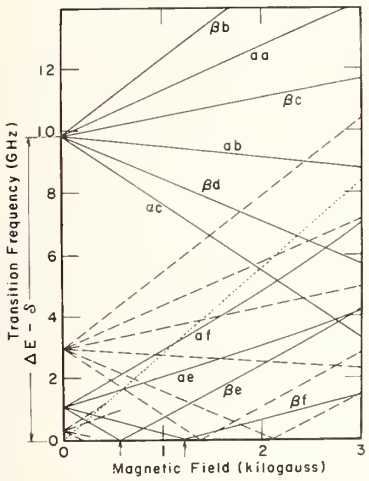


FIGURE 3. *Plot of transition frequencies versus magnetic field for the* n=2 *levels (solid lines) and some of the* n=3 *levels (broken lines) of hydrogen.*
The dotted line indicates the position of the electron cyclotron resonance transition.

forwardly related to the zero-field interval $\Delta E - S$. It was found that the three β-state transitions, βb, βc, and βd were not good candidates for precision measurements, for the following reason. As indicated by the two "zero-frequency" points for the βe and βf transitions, the metastable β-state is prone to undergo transitions in these regions if *static* electric fields are present, an effect exploited in the level-crossing measurements of Robiscoe and Cosens [3]. In our apparatus, such β-state quenching effects are found to distort the high frequency resonance transitions involving the β-state, and thereby

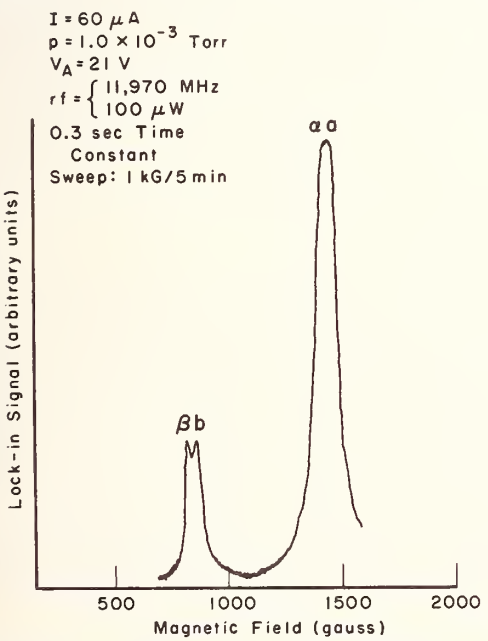


FIGURE 4. *Chart recording of the αa and βb resonances at a microwave frequency of 11970 MHz.*

render them unsatisfactory for precision measurements. The metastable α-state, on the other hand, does not suffer a similar quenching effect, except at rather higher magnetic fields. The transition αa was selected for measurement, and later transitions αb and αc were also measured. Figure 4 shows a typical chart recording of the αa transition, taken as a function of magnetic field at a microwave frequency of 11970 MHz. The closest transition that might cause overlap problems is βb, some 600 G lower.

Precision data taking consists of signal measurements at five selected magnetic field values, located symmetrically astride the resonance curve. Data analysis involves a series of procedures, which may be listed as follows:

(i) a least squares adjustment of the data is made to a theoretical function describing the resonance line shape; this adjustment produces a value for the line center;
(ii) this result is compared with that computed from energy levels calculated by diagonalization of the Hamiltonian for the $n=2$ levels of hydrogen; any difference between these two may be removed by introducing a correction Δ to the value of $\Delta E - S$ used in diagonalizing the Hamiltonian;
(iii) the correction, Δ, is examined for dependence on experimental parameters, in particular the rf power level, the electron beam current, and the gas pressure; a linear extrapolation of the Δ's is used to remove such residual dependence as is found;
(iv) the extrapolated value of Δ is applied to yield the final experimental result for $\Delta E - S$.

Of special interest in this experiment is the description of the observed resonance line shape. We developed a model to account for the quenching of the metastable hydrogen atoms by rf and other quenching mechanisms.

We denote by μ the probability per unit time that a metastable atom is quenched by the rf radiation, and let λ be the corresponding probability for quenching by any other process. Then suppose metastable atoms, each with velocity v, are excited at a rate r at a point P inside the waveguide. In the steady state, the rate of arrival of unquenched atoms at a distance L from P is

$$r \exp(-(\mu+\lambda)L/v).$$

The steady-state flux of light leaving this path, being proportional to the total number of metastables quenched per unit time along the path, is given by the expression

$$F = r[1-\exp(-(\mu+\lambda)L/v)] \qquad (1)$$

to within a proportionality constant. When the rf is off, $\mu=0$. Hence, the signal, which is proportional to the change in the light flux on turning the rf on, may be written

$$F_0 = r\cdot\exp(-\lambda L/v)\ (1-\exp(-\mu L/v)). \qquad (2)$$

TABLE 1. *Summary of results and corrections*
The αc results were not incorporated into the final average, but are included for completeness. See text for details.

Transition	αa	αb	αc
Number of runs	148	62	71
Frequency ---------- MHz	11970	9170	7430
Magnetic field at line center ---------- gauss	1465	1860	1090
Extrapolation coefficients:			
Pressure term ---------- kHz millitorr^{-1}	+89.2(1.72)	−12.9(1.9)	+29.5(105.)
Current term ---------- kHz mA^{-1}	+335(284)	−237(77)	+3.54(22)×10^3
Raw extrapolated result ($\Delta E-S$) ---------- MHz	9911.402(18)	9911.394(10)	9911.027(050)
Corrections and uncertainties (MHz):			
NMR calibration	−0.058(12)	+0.015(03)	+0.039(008)
Relativistic energy	+0.019	−0.002	−0.012
Quadratic Zeeman energy	+0.001	+0.006	+0.003
Stray fields ($\leq$2V/cm)	−0.001(02)	−0.006(06)	0.000
$\mathbf{v}\times\mathbf{H}$ Stark shift	(20)	(31)	(001)
Overlapping resonance	(10)	(30)	(300)
Computer roundoff	(02)	(02)	(002)
Final results ---------- MHz	9911.363(31)	9911.407(45)	9911.057(304)

Weighted average $\Delta E-S$ ---------- MHz ---------- 9911.377(26)

The quenching rate μ has been shown by Lamb [4] to be given by

$$\mu_{nm}=\frac{\gamma\,|\,\langle m|\,e\mathbf{E}_0\cdot\mathbf{r}\,|\,n\rangle\,|^2}{4\hbar^2[(\omega_{nm}-\nu)^2+1/4\gamma^2]}. \qquad (3)$$

This relates to the case of two levels n and m (S and P levels respectively) differing in energy by $\hbar\omega_{nm}$ and coupled by an alternating electric field of amplitude E_0 and circular frequency ν. The decay rate of the short-lived P state is denoted by γ. The numerator of the above expression contains the modulus squared of the electric dipole matrix-element between the pair of levels m and n. To allow for rf quenching transitions between other states, the expression for μ_{nm} in eq (3) is summed over all allowed values of m, to give

$$\mu_n=\sum_{m\neq n}\mu_{nm}. \qquad (4)$$

It remains to average the signal expression, eq (2), to allow for the distribution in velocity, variation of rf field within the waveguide, and variation in path lengths and starting points of metastable atoms produced by the electron beam. The averaging may be carried out in a formal manner, sufficient to yield a functional expression for the quenching though involving unknown coefficients. However, a comparison with experiment serves to determine these coefficients, and an excellent working quenching function is thereby obtained. This is illustrated in figure 5, in which the signal amplitude at the center of the αa resonance is plotted against the rf power. The experimental points are seen to be fitted very well by the theoretical curve.

Another source of concern in the present experiment is the possible variation of r, the rate of excitation of metastable atoms, with magnetic field. This could introduce a serious asymmetry into the resonance curve. The effect was finally disposed of by the following normalization procedure. The signal was everywhere measured twice, once using a low rf power level, and again with a sufficiently high rf level to broaden the resonance considerably, while by no means saturating it. By taking the ratio of the two signals, the factor r cancels out. All precision data was taken using this rf normalization method, and analyzed accordingly.

A discussion of other sources of systematic error, and of several small, calculable corrections to the data, must here be foregone for lack of space. Such matters as the Stark effect, collisions, Doppler shift, cascade effects, and angular distribution of the

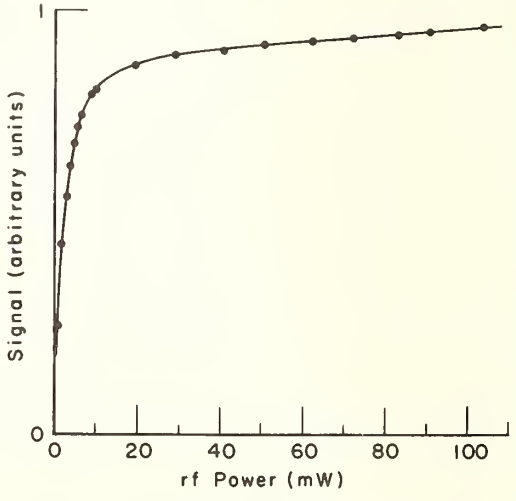


FIGURE 5. *Plot of signal amplitude, at the center of the* αa *transition, versus rf power.*
A theoretical curve has been fitted to the experimental points.

Lyman-α light, among others, are treated in a forthcoming paper [5].

We present next a summary of the experimental results, given in table 1. This shows most of the significant information about the three measured transitions, αa, αb, and αc. The results were found to depend slightly on gas pressure and bombardment current, though not on rf power level. The coefficients of a least squares fit to a linear extrapolation formula in both pressure and current are tabulated. A sizeable number of experimental runs were taken under conditions where the extrapolation to zero pressure and current shifted the results by only 20 to 50 KHz. The extrapolated results for $\Delta E - \mathcal{S}$ are shown in the table, along with the standard deviations of the mean. It was found convenient to incorporate a number of small corrections in an additive fashion to the results at this final stage, although strictly they should have been included in the data analysis at the outset. Because of their small size, there is no loss of precision to the present level of accuracy. The NMR calibration is the correction for the small magnetic field difference between the experimental region and the NMR probe. The relativistic energy and quadratic Zeeman energy corrections allow for two terms which were omitted from the working Hamiltonian. The correction for stray fields allows for electric fields which may be present when the beam current is zero. Also listed are other sources of uncertainty attributable to the motional Stark effect, overlapping resonances, and computer roundoff. The uncertainty figures in this part of the table are not of a statistical nature, although an attempt has been made to give realistic 68 percent confidence intervals. When these corrections are included, and all uncertainties summed in quadrature, final results are obtained, as shown in table 1. It is seen that the result for αc has a much greater error bar than the other two. This is due primarily to uncertainty in the correction for overlap from the nearby stronger resonance, βd. In view of this uncertainty, the αc result was not included in the final weighted average, which was based only on the αa and αb measurements, with the result

$$\Delta E - \mathcal{S} = 9911.377(26) \text{ MHz}.$$

3. Discussion

This result may be compared with other recent experimental measurements of the same fine structure interval [6]. It may also be combined, in a variety of ways, with other existing measurements of $\mathcal{S}$ and ΔE, so as to provide either a value for the fine structure constant, α, or a value for $\mathcal{S}$. Of the many possible alternatives, we here select but one. A theoretical value for $\Delta E - \mathcal{S}$ has been given by Appelquist and Brodsky [7], viz., $\Delta E - \mathcal{S} = 9911.12 \pm 0.22$ MHz. This incorporates their recalculation of the fourth order radiative correction to the Lamb shift, $\mathcal{S}$, and also embodies the value $\alpha^{-1} = 137.03608(26)$ recommended by Taylor, Parker, and Langenberg [1]. The error bar on the theoretical value for $\Delta E - \mathcal{S}$ is regarded as a limit of error, whereas the experimental result is quoted with a one standard deviation uncertainty. It is apparent that the agreement is poor, and this remains a source of concern to us.

4. References

[1] See, for example, the review article by Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).

[2] Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **79**, 549 (1950).

[3] Robiscoe, R. T., Phys. Rev. **168**, 4 (1968); Cosens, B. L., Phys. Rev. **173**, 49 (1968).

[4] Lamb, W. E., Jr., Phys. Rev. **85**, 259 (1952).

[5] Kaufman, S. L., Lamb, W. E., Jr., Lea, K. R., and Leventhal, M., to be submitted to Physical Review. See also Phys. Rev. Letters **22**, 507 (1969).

[6] Shyn, T. W., Williams, W. L., Robiscoe, R. T., and Rebane, T., Phys. Rev. Letters **22**, 1273 (1969); Vorburger, T. V., and Cosens, B. L., Phys. Rev. Letters **23**, 1273 (1969). See also papers by these authors in these Proceedings.

[7] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970). See also invited paper by S. J. Brodsky, these Proceedings.

DISCUSSION

W. E. Lamb, Jr.: This discrepancy between theory and experiment is very worrisome to us. I should point out two things. This is a very different sort of experiment. It's a "dirty" experiment. The work we heard about earlier is concerned with signals of the order of 10^4 or 10^5 metastable atoms per second. In other words, in the apparatus there is one metastable atom, or less, at any one time. In the work that Dr. Lea has described there is a very much larger signal. It's possible to take resonance curves very, very fast. The only trouble is we don't know what they mean. I think myself the two weakest points are in the possible effects of stray electric fields in a rather unclean environment and possibly the effects of complications from higher states that we have somehow not considered because we are dealing with metastable atoms that are formed by the dissociation of the hydrogen molecule by electron bombardment and do not have the purifying effects of a metastable beam to get rid of highly excited states of the atom or molecule.

Atomic Beams Measurements of the Fine Structure of H, $n=2$

R. T. Robiscoe

Department of Physics, Montana State University, Bozeman, Mont. 59715

We discuss the most recent experimental systematic corrections to atomic beams measurements of the fine structure of the first excited state of atomic hydrogen, paying particular attention to those measurements of the Lamb Shift S with a claimed accuracy of 0.1 MHz. The newest correction has to do with kinematic effects on the beam velocity distribution. Applying this correction to the so-called level crossing experiments, we find the best expurgated value of S to be 1057.90 ± 0.10 MHz. We compare this result with the most recent theoretical value, and comment on its reliability.

Key words: Experimental corrections; Lamb shift; quantum electrodynamics.

There have been at least seven high precision experiments done to measure some aspect of the fine structure of the $n=2$ level of atomic hydrogen [1–7]. Five assorted values of the fine structure constant α have been quoted, to within 10 ppm, as the results of these measurements [2, 4, 5–7]. With the exception of the optical level crossing measurement [4], these α values depend upon a *separate* knowledge of the $2S_{1/2}$–$2P_{1/2}$ level separation S (the Lamb shift) and the $2P_{3/2}$–$2S_{1/2}$ level separation $\Delta E - S$. These two intervals are added together to give the total fine structure interval ΔE, from which a value of α is calculated. Thus, quite apart from its intrinsic interest as a quantum electrodynamical quantity, the value of S plays a control role in the determination of α from hydrogen fine structure.

While $\Delta E - S$ has been measured four times [2, 5–7], we have only two measurements of S itself [1, 3]. Both S measurements were done by an atomic beams resonance method to an accuracy in S of 0.1 MHz, or about 100 ppm. This is the accuracy necessary for an α determination to 10 ppm, and it required a resolution of the resonance line observed to within 0.1 percent of its full width at half maximum. At this extremely high level of resolution, there arise many systematic corrections to the observed line centers. Indeed, both S measurements had to be corrected for systematic effects by several times the claimed accuracy of 0.1 MHz. Altogether, then, the hydrogen fine structure value of α depends to a large extent on the accuracy of the systematic corrections made in the two S measurements by the atomic beams method.

In this talk, I wish to discuss the most recent experimental systematic corrections applied to the atomic beams measurements of hydrogen fine structure. These corrections relate to the assumed velocity distribution of the $2S_{1/2}$ metastable atomic beams used in the experiments, and are mainly of importance in correcting the S values; they are negligible in the $\Delta E - S$ work. Unfortunately, the beam velocity distributions were never directly measured during the S experiments, so that the corrections I will discuss are necessarily *a posteriori*. Nevertheless, they lead to a satisfactory agreement between theory and experiment for S, and do not disturb the reasonable consistency among hydrogen fine structure α values or their agreement with the α value from the ac Josephson work [8].

To show why the metastable beam velocity distribution is important in the S experiments, we must briefly describe the experimental method. In both the Lamb [1] and level crossing (LC) [3] experiments, a beam of metastable 2S atoms is sent at velocity v into a transition region where, in the presence of a magnetic field H, an applied rf electric field E induces $2S-2P$ transitions. E is kept fixed while H is swept through some portion of the Zeeman diagram to trace out a $2S-2P$ resonance. By virtue of passing through H, the 2S atoms experience a motional electric field of magnitude

$$E_m = (v/c)\,\mathrm{H}\sin\theta, \tag{1}$$

where $\theta = \angle(\mathrm{H}, \text{beam trajectory})$. In the Lamb experiments, where $\theta = 90°$ (except for beam divergence), E_m was of order 5 V/cm, while in the LC experiments, where $\theta \simeq 0°$ (except for beam divergence and a small misalignment between the field and beam axes), E_m was of order 0.1 V/cm. Clearly, any effects of E_m must be averaged over the beam velocity distribution, which is usually assumed to be Maxwellian.

There are two systematic corrections which depend upon E_m and, thus, upon the beam velocity distribution. First, there is a relative Stark shift of the 2S and 2P energy levels due to E_m. For the Lamb experiments, this amounted to a correction of 0.1 to 0.3 MHz in S [9], while for the LC experiments, it was negligibly small. The second correction, which I call the motional field (MF) asymmetry, is more subtle. E_m scales as $v\,\mathrm{H}$, so that it preferentially quenches high velocity metastables and reduces the

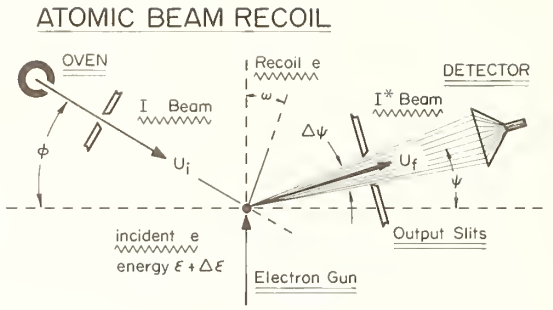


FIGURE 1. *Experimental schematic.*
A beam of ground state particles I effuses from an oven and is incident on a beam of electrons. The electron impact excites metastable particles I* which exit toward a detector in the shape of a "fan." Generally, the angles are adjusted so as to provide maximum I* signal.

average velocity of the surviving beam. This effect becomes more pronounced as H is increased. Thus, as we scan a resonance, the average velocity of the beam decreases, and the surviving atoms spend more and more time in the transition region. Since the number of metastables quenched is proportional to the transit time, the upper wing of the resonance line is raised relative to the lower. The apparent line center shifts upward, which results in an erroneous value for S. For the Lamb experiments, the correction for this effect amounted to about 0.1 MHz in S, while for the LC experiments, it was about 0.2 MHz.

As we have said, both the Stark shift and MF asymmetry corrections depend upon the metastable velocity distribution. In both the Lamb and LC experiments, and for the corrections cited above, the distribution was assumed to be $U^2 \exp(-U^2)$, a form which follows from simple arguments based on kinetic theory and a consideration of the beam excitation conditions. This form was indirectly verified in the Lamb experiments by measuring the inverse first moment of the distribution [10]. In the LC experiments, this form was assumed *a priori*, and no attempt was made to measure it. Moreover, in the initial report of the LC S value [3], it was implicitly assumed that motional fields as small as 0.1 V/cm were too small to substantially distort the velocity distribution in a manner leading to the MF asymmetry; consequently, that correction was ignored. The Lamb value for S and the initial LC value disagreed by 0.3 MHz, which was about three times the upper limit of error assigned to either experiment. It was only later I realized that even very small motional fields could substantially distort the velocity distribution in the neighborhood of a crossing point. An *a posteriori* calculation of the MF asymmetry correction reduced the LC S value by about 0.2 MHz and reconciled it with the Lamb value (to within 0.1 MHz) [11].

While the experimental S values were reconciled, there was still a discrepancy with theory of about 0.3 ± 0.2 MHz. Since the error here was the simple sum of the upper limits of error associated with both the calculation and measurement (about 0.1 MHz for each), then the discrepancy had to be taken seriously. If the theory was right, then the hydrogen fine structure α value could be shifted by as much as 14 ppm, well outside its quoted precision. If the experiment was correct, then there had to be something wrong with the QED calculations, work which had been done and redone by a number of people. Amazingly enough, simply by careful analysis of the internal consistency of the various existing hydrogen fine structure measurements, Taylor, Parker, and Langenberg had concluded their 1969 review article on fundamental constants with the statement "that the disagreement between S(exptl) and S(theory) is probably real and that the theory of the Lamb shift needs revision." This statement has turned out to be correct, and that justifies in some measure the sort of work done by those who review the fundamental constants. Nevertheless, with some residual faith in the theory, and an awareness of the many systematic corrections to the S measurements, the natural place to look for an error was in the S experiments. And there, the obvious thing to look at was the beam velocity distribution, which had never been directly measured.

Figure 1 is a schematic diagram of what happens to the hydrogen beam during the excitation process employed in the S experiments. Ground state atoms are incident on an electron beam which is run at an energy just above threshold for excitation to the metastable state. Depending on their incident velocity, the resulting metastables are spread out into a "fan" which moves toward the detector. With a velocity distribution $U^3 \exp(-U^2)$ of ground state atoms out of the oven, and a transit time factor $1/U$ for the excitation process, the metastable

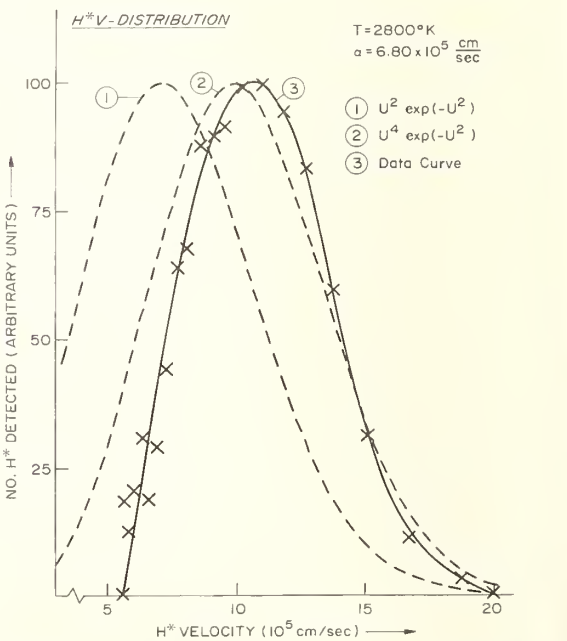


FIGURE 2. *Measured* H* *velocity distribution.*
The experimental points are compared with a theoretical curve (solid line) and with curves of the form $U^n \exp(-U^2)$ for n=2 and 4.

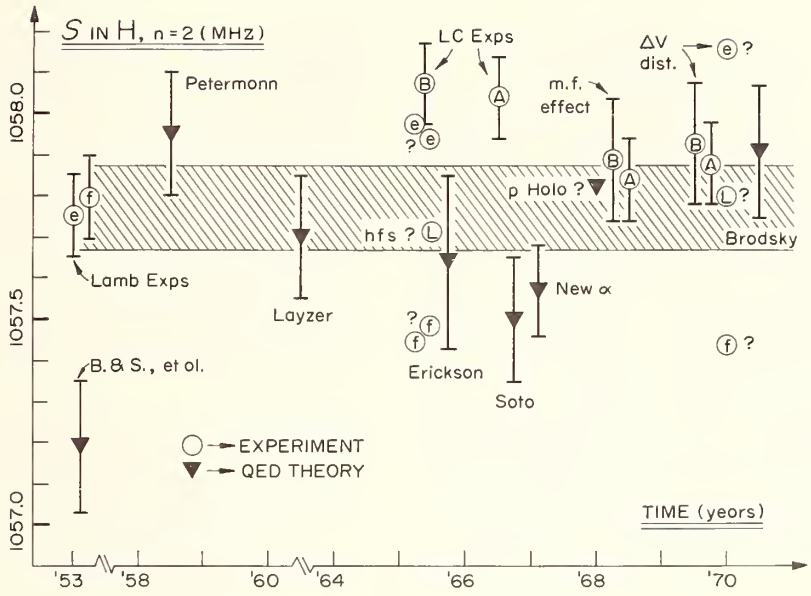


FIGURE 3. *History of values of the Lamb shift* S *in* H, n=2.

e and *f* refer to the values measured by Lamb and collaborators from transitions αe and αf (ref. [1]). *A* and *B* refer to the values measured at the 538 G and 605 G level crossings (ref. [3]). L refers to the mean of the Lamb values provisionally corrected first for neglect of a hyperfine structure (hfs) term (noted in Phys. Rev. **138A**, 22 [1965]), and then for the new velocity distribution (ref. [14]). The level crossing values have been corrected for a motional field (m.f.) effect (ref. [11]), and for the new velocity distribution (ref. [14]). The latest theoretical result is that of Appelquist and Brodsky (ref. [15]). References to the earlier theoretical work can be found in the paper by Taylor, et al. (ref. [8]).

distribution is $U^2\exp(-U^2)$ downstream. But this is only true if *all* metastables are collected. If there is any collimation of the downstream beam, say by finite size of the detector, then both the very fast and very slow metastables, which occur on the extreme edges of the "fan," are lost to the beam. This distorts the $U^2\exp(-U^2)$ distribution in a way which is quite sensitively dependent on the beam alignment and apparatus geometry.

By solving the momentum and energy conservation equations for the situation in figure 1, it is possible to calculate what the distribution will be for a given geometry. One predicts upper and lower cutoff velocities, and an overall distribution which is considerably faster in typical cases than the usually assumed $U^2\exp(-U^2)$ distribution. For example, at the excitation threshold, and for a beam alignment producing maximum detected signal, the most probable beam velocity is $U_{mp}=\sqrt{2}$ for the actual distribution, while it is $U_{mp}=1$ for a $U^2\exp(-U^2)$ distribution. It is somewhat surprising that the results of this straightforward calculation are not better known, and have not (to my knowledge) become part of the literature on atomic and molecular beams.

Figure 2 shows an experimental confirmation of the calculation. It is the velocity distribution of a beam of metastable hydrogen atoms actually used in the latest hydrogen fine structure measurements. The apparatus was aligned to give maximum metastable signal, and the metastable time of flight was measured as the electron gun was pulsed. The observed peak and cutoff velocities agree with absolute calculated values to within 10 percent. The experimental points are compared with a theoretical curve (solid line) [12], and with curves of the form $U^n\exp(-U^2)$ for $n=2$ and 4. Clearly, the observed distribution is much faster than $n=2$ (with $U_{mp}=1$); it is more nearly $n=4$ (with $U_{mp}=\sqrt{2}$). The deficiency of slow metastables is particularly marked. Such deficiencies have been noted in other experiments with light metastable atoms, and are usually blamed on poor definition of oven temperature or self-scattering in the beam. We believe the present kinematic effects, due to recoil of the excited atoms and collimation of the beam, are in fact a major cause in distorting the velocity distribution. This fact is quite interesting in itself, and the kinematic effects are presently being studied [13].

Although the beam velocity distributions were not directly measured during the S experiments, clearly the indication is that they were considerably faster than $n=2$. This means that the motional field corrections are changed. For purposes of calculation, it is sufficient to assume an $n=4$ distribution. Correcting the level crossing work is unambiguous, as it was done on an apparatus virtually identical to that used to measure the distribution of figure 2. The Stark shift correction remains negligible, while the motional field asymmetry correction is changed by 0.04 MHz. The corrected mean value from the level crossing work becomes

$$S(\text{LC expt}) = 1057.90 \pm 0.10 \text{ MHz}. \qquad (2)$$

This value has already been reported [14]. The total systematic corrections amount to about 0.2 MHz, and are good to about 10 percent. The quoted precision is at least twice the standard deviation of the

mean of the means of ten independent runs comprising over 200 line center measurements.

Correcting the Lamb experiments for the kinematic distortions to the beam velocity distribution is less certain. The apparatus geometry was rather different than in the level crossing work, and beam alignment and collimation were less critical. Measurements of the first inverse moment of the distribution were consistent with a form $U^n \exp(-U^2)$ for $n=2.5$ [10]. However, the analysis used is rather sensitive to the assumed velocity cutoffs. For example, for the same data, $n=3.6$ if the velocity range was $0 \leq U \leq 1.5$ rather than $0 \leq U \leq \infty$. Unless the beam geometry during this measurement was identical (in alignment and collimation) to that used during the S measurements, it is possible that velocity selection effects were introduced inadvertently. If they were as large as to require using an $n=4$ rather than an $n=2$ distribution, then the Stark shift corrections would be increased by a factor of about 5/3, and the mean value of S from the Lamb experiments would be raised by about 0.09 MHz. But since there is no direct evidence to support this correction, we believe it should be considered at most as a possible source of previously unaccounted systematic error.

To conclude, we show a comparison of theory and experiment in figure 3. Here, there is either a very long story or a very short one, depending on how strongly you feel about shifts of a few tenths of an MHz in the value of S. The standard of comparison for all S values has been, and still is, the original numbers measured by Lamb and collaborators. The initial LC values of S are shown (ca. 1966) in disagreement with the Lamb S. By 1968, the experimental values had been reconciled by discovery of the motional field asymmetry correction in the LC work. The new velocity distribution correction moves the LC values slightly upward. If one uses an $n=4$ distribution for the Lamb experiments, the Lamb values are moved apart by about 0.7 MHz, but the mean is changed only slightly. The most notable recent change has been in the theory, namely the 0.4 MHz shift from the Soto value to the new value quoted by Brodsky and Appelquist [15]. At present, there is good agreement between both experiments and between theory and experiment. This seems to be a fortunate occurrence, just now at the end of the latest review of hydrogen fine structure, and at a time when interest in the intrinsic accuracy of those measurements regarding values of α has shifted to the ac Josephson effect experiments.

References

[1] Triebwasser, S., Dayhoff, E. S., and Lamb, W. E., Jr., Phys. Rev. **89**, 98 (1953).
[2] Dayhoff, E. S., Triebwasser, S., and Lamb, W. E., Jr., Phys. Rev. **89**, 106 (1953).
[3] Robiscoe, R. T., and Cosens, B. L., Phys. Rev. Letters **17**, 69 (1966).
[4] Metcalf, H., Brandenberger, J. R., and Baird, J. C., Phys. Rev. Letters **21**, 165 (1968).
[5] Kaufman, S. L., Lamb, W. E., Jr., Lea, K. R., and Leventhal, M., Phys. Rev. Letters **22**, 507 (1969).
[6] Shyn, T. W., Williams, W. L., Robiscoe, R. T., and Rebane, T., Phys. Rev. Letters **22**, 1273 (1969).
[7] Cosens, B. L., and Vorburger, T. V., Phys. Rev. Letters **23**, 1273 (1969).
[8] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Revs. Mod. Phys. **41**, 375 (1969).
[9] See Table XX of reference 1.
[10] See Appendix VII of reference 1.
[11] Robiscoe, R. T., Phys. Rev. **168**, 4 (1968).
[12] The theoretical curve was worked out by T. Rebane, The University of Michigan, Ph.D. thesis, 1970, unpublished.
[13] Pearl, J. C., Donnelly, D. P., and Zorn, J. C., Phys. Letters **30A**, 145 (1969).
[14] Robiscoe, R. T., and Shyn, T. W., Phys. Rev. Letters **24**, 559 (1970).
[15] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970).

DISCUSSION

T. V. Vorburger: Is it possible that conditions in the source—for instance, the Knudsen number—could have anything to do with changing the original ground state velocity distribution?

R. T. Robiscoe: I think that's very possible. But what we discovered from that one graph I showed you where there was a theoretical fit is that the purely kinematic effects seem to explain the velocity distribution. The conditions inside the oven I think are probably a great mystery, and indeed if there is a jet flow out of the oven, if there is a non-equilibrium situation in the oven, one would expect temperature-dependent effects. But as far as we can see, the major distortion in the velocity distribution, the move, the 40 percent increase in the mean speed of these particles, is mainly due to these kinematic and collimation effects.

S. J. Brodsky: I should mention there's a lot of people involved in the theory, and maybe just for the record I could mention some of the people who made major contributions to the theory of the Lamb shift. Bethe and Lamb; Karplus, Klein, and Schwinger; Baranger, Bethe, and Feynman; Layzer; Kallen and Sabry; Erickson and Yennie; Fried and Yennie; Salpeter; Fulton and Martin; Grotch and Yennie.

Separated Oscillatory Field Measurement of the Lamb Shift*

C. W. Fabjan and F. M. Pipkin

Lyman Laboratory, Harvard University, Cambridge, Mass. 02138

A new technique of observing the Lamb shift, which combines beam-foil and radiofrequency spectroscopy techniques, is described. Measurements are made on fast H-atoms in zero magnetic field by sweeping the frequency of the rf-field. It is described how this technique allows the use of Ramsey-separated oscillatory rf-fields with a resulting decrease in the observed line width of the resonance. A rf hyperfine state selector is used to reduce the effect of overlapping lines.

Key words: Beam foil spectroscopy; hydrogen; Lamb shift; line narrowing; separated oscillatory fields.

1. Introduction

Precise measurements of the Lamb shift of hydrogenic atoms are among the most sensitive tests of quantum electrodynamics [1]. The precision of the previously performed measurements was limited by the natural line width of the resonances and by the uncertainty of the corrections inherent in the experimental techniques used [2].

In this paper we wish to report a new method of measurement, which combines beam foil spectroscopy [3] and radiofrequency spectroscopy. The Lamb shift resonances are observed in a fast H-beam, which permits the use of separated rf fields with a resulting decrease in the linewidth of the resonances [4, 5].

2. Discussion of the Experimental Technique

2.1. Production and Observation of the H-Beam

Figure 1a shows a schematic of the apparatus used. Protons are extracted from an rf-ion source, accelerated and focused into a parallel beam. The proton energy was kept in the range from 20 keV to 40 keV, which corresponds to a velocity range of 2.0 to 2.8×10^8 cm/s. At these energies about 90 percent of the protons capture an electron during their passage through the carbon foil and form excited or ground state H-atoms. Subsequently transitions between the $S_{1/2}$ and $P_{1/2}$ states are induced in the rf-interaction region (RF1 chamber, RF2 chamber). Since the P-states decay much more rapidly, the $S_{1/2}$ state population will be reduced; this is detected downstream as a decrease in the count rate of the photomultiplier PM. Observations were made on the $n=3$ state by using an interference filter to select Balmer-α-photons. We kept the proton intensity below 5×10^{12} atoms/s to increase the lifetime of the foil. This intensity resulted in a count rate at the PM of typically 70,000 counts/s.

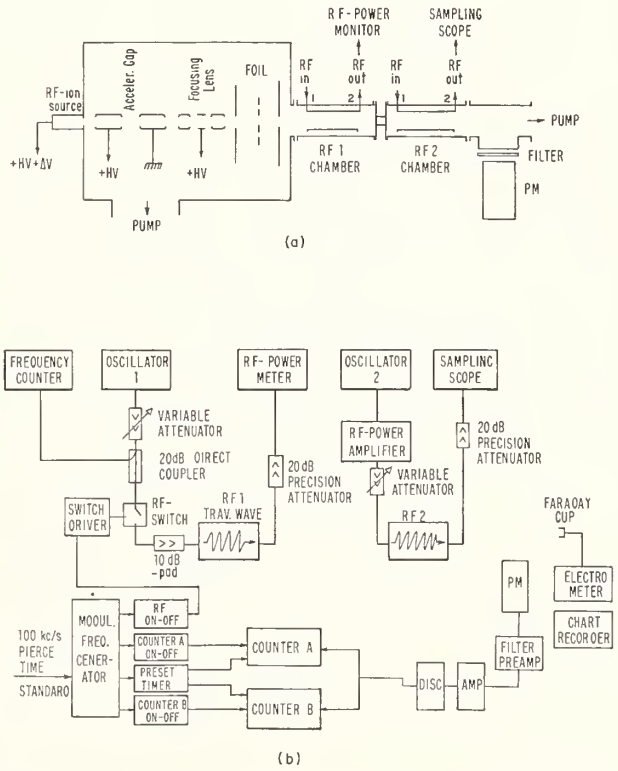


FIGURE 1. (a) *Schematic diagram of experimental apparatus.* (b) *Block diagram of electronics used to observe the Lamb shift with one rf-spectroscopy field.*

* Research supported in part by NSF grants GP-8551 and GP-22787 and in part by a grant from the Milton Fund.

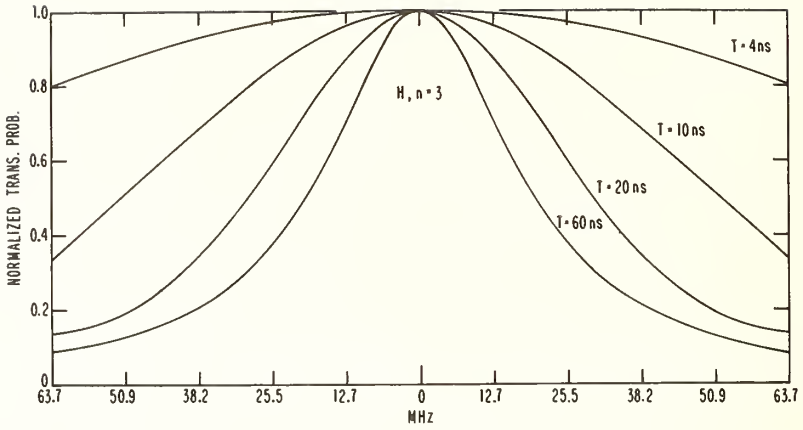


FIGURE 2. *Transition probability as a function of frequency and time* T *spent in the rf field computed with eq (5).*

2.2. Measurement of Resonances

In all previous experiments the atoms were studied in the presence of an external magnetic field. Transitions were induced at a fixed frequency of the rf-field and they were swept through resonance by changing the magnetic field. This procedure introduced a number of sizeable corrections, which were not always easy to understand [2]. For this reason we decided to study the atoms in zero magnetic field and to sweep the transitions by varying the frequency of the rf field.

The following approach was used to keep the rf power constant over the frequency range of interest. A coaxial transmission line was brought into the chamber and connected to two parallel plates between which the beam traveled. The other ends of the rf plates were connected to a thermistor, which acted as a 50 Ω-termination to very high accuracy and was used to monitor the rf power. Thus, a traveling wave was set up and ideally the total power incident on the entrance port was propagated along the rf plates and absorbed by the thermistor. The VSWR at the rf ports and the insertion loss of the rf plates was measured as a function of frequency to correct for any mismatch. The rf power correction was at most 2 percent over a bandwidth of 100 MHz.

2.3. Measurements With One RF Spectroscopy Field

The $S_{1/2}$–$P_{1/2}$ transition probability as a function of frequency and amplitude of the rf field and the time t spent in the rf region can be computed by solving the time dependent Schrödinger equation. If the wavefunction of the system is written in the form

$$\psi = c_s(t) \mid S\rangle + c_p(t) \mid P\rangle \tag{1}$$

one obtains a system of two coupled differential equations:

$$\dot{c}_s = i\omega_{sp} c_p \exp[-i(\omega-\omega_s+\omega_p)t] - \tfrac{1}{2}\gamma_s c_s \tag{2}$$

$$\dot{c}_p = -i\omega_{sp}^* c_s \exp[-i(\omega+\omega_s-\omega_p)t] - \tfrac{1}{2}\gamma_p c_p \tag{3}$$

where

$$\omega_{sp} = -(eE_0/2\hbar)\langle S \mid \mathbf{r}\cdot\mathbf{d} \mid P\rangle. \tag{4}$$

If one assumes for the initial conditions

$$c_s(0) = 1 \qquad c_p(0) = 0$$

the solution is found to be

$$c_s(t) = (A_- - A_+)^{-1} \cdot \exp(-\gamma_s/2 - i\omega_s)t \cdot [A_- \exp A_+ t - A_+ \exp A_- t] \tag{5}$$

where

$$A_\pm = \tfrac{1}{2}[Q/2 + i(\omega-\omega_s+\omega_p)] \pm \{[Q/2+i(\omega-\omega_s+\omega_p)]^2 - 4 \mid \omega_{sp} \mid^2\}^{1/2} \tag{6}$$

$$Q = \gamma_p - \gamma_s. \tag{7}$$

It is interesting to note that apart from a factor $\exp(-\gamma_s t/2)$ the solution depends only on the difference $\gamma_p - \gamma_s$. Figure 2 shows typical line shapes.

In hydrogen the analysis is complicated due to the presence of hyperfine interactions. As indicated in the level diagram, figure 3, three transitions are possible

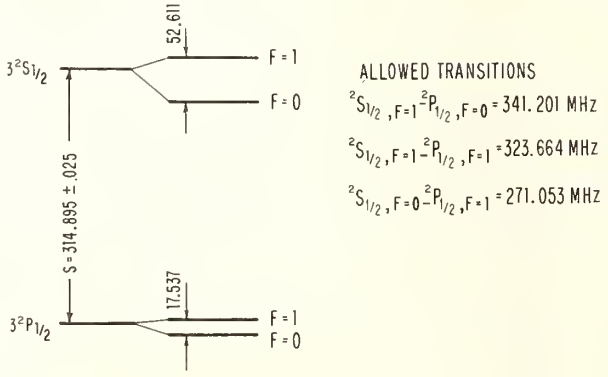


FIGURE 3. *Level diagram of the* $S_{1/2}$ *and* $P_{1/2}$ *states in* H, n=3 *in zero magnetic field.*
Quoted intervals are given in units of MHz.

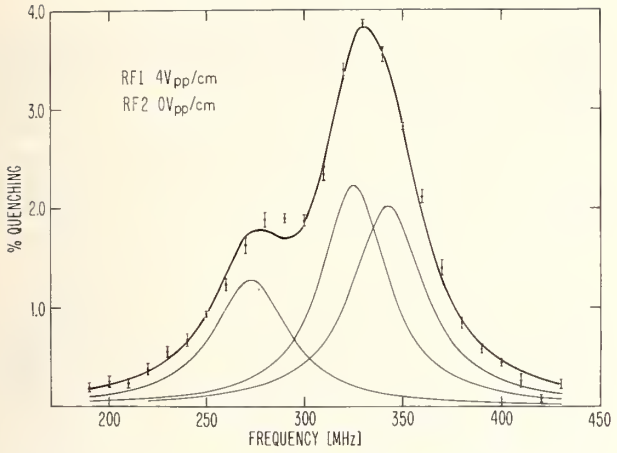


FIGURE 4. *Typical line scan of Lamb shift resonance observed with one rf-spectroscopy field.*

in zero magnetic field. These can clearly be seen in a typical line scan, figure 4.

Also shown is a computer least squares fit to three resonances. To obtain this fit the hyperfine spacing of these three resonances was assumed to be known, but the Lamb shift and the widths and heights of the individual components were adjustable parameters.

To allow the measurement of a single transition we investigated the use of a rf hyperfine state selector. As indicated in figure 1a, the H-beam first traverses the spectroscopy region RF1 and subsequently enters region RF2, in which an rf field corresponding to the transition frequency $S_{1/2,F=1}-P_{1/2,F=1}=323.7$ MHz or $S_{1/2,F=1}-P_{1/2,F=0}=341.2$ MHz is applied. This field reduces the number of $S_{1/2,F=1}$ states and hence the PM registers predominantly light from the $S_{1/2,F=0}$ state. A typical scan employing the hf-state selector is shown in figure 5a. The relative contribution of the $S_{1/2,F=1}$ states has been reduced considerably.

Figure 1b shows the block diagram of the electronics used for these measurements. Two scalers were used to compare the number of Balmer-α-protons emitted when the rf field was alternately switched on and off.

2.4. Two-Oscillatory-Field-Measurements

It was shown by Ramsey [4] that the linewidth of a resonance transition between two stable levels can be reduced, if instead of one rf field with length L two short coherently oscillating rf fields, which are separated by a distance L, are used. A similar advantage can be gained, if this type of spectroscopy is performed on decaying levels [5].

In addition to this interference-line narrowing the method of two coherent rf fields provides a very convenient and efficient way of selecting P-states which have lived a time $t>\tau_p$, the average lifetime of the P-state. This time t is an adjustable experimental parameter and is to a large extent determined

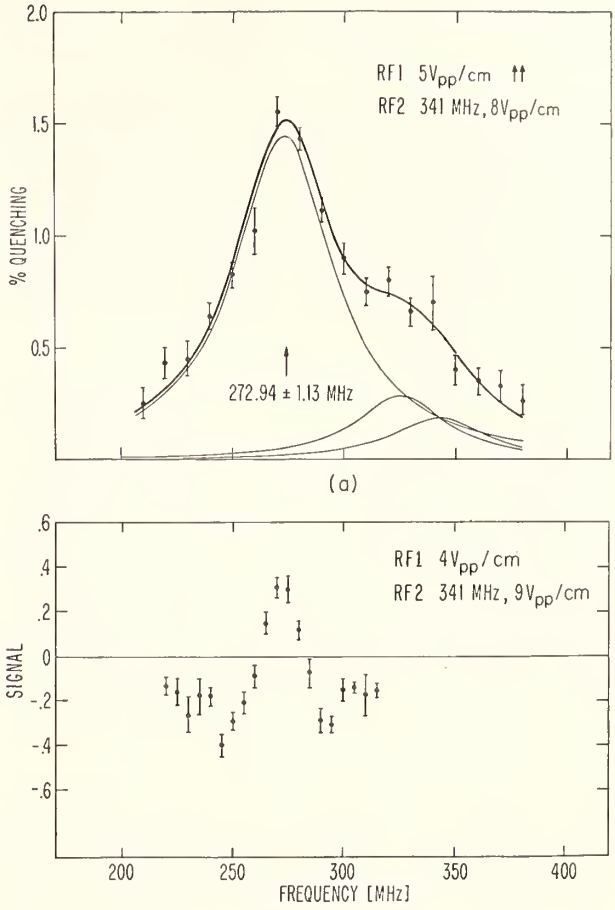


FIGURE 5. (a) *Scan of Lamb shift resonance with one rf-spectroscopy field. The rf-hyperfine state selector was set at the frequency of the* $S_{1/2,F=1}-P_{1/2,F=0}$ *transition.* (b) *The Lamb shift transition observed with two coherently oscillating fields with* $\tau=12$ *ns and* $T=7$ *ns. The same rf-quenching field as in figure 4 was employed. The signal is the difference between the fractional changes in the count rate, when the relative phase of the two rf fields is switched from 0° to 180°.*

by the flight time T of the atoms between the separated rf fields. Thus a further "uncertainty principle"-line narrowing is possible. Extending the treatment given in [5] one finds for the S-state amplitude of an atom, which traveled through the first rf field for τ seconds, through a field free region for T seconds, and again through a rf field for time τ, the following expression:

$$c_s(\tau+T+\tau)=(A_--A_+)^{-2}[\exp(-\gamma_s/2)\,(\tau/2+T)\,]\cdot\{[A_-\exp A_+\tau-A_+\exp A_-\tau]^2-\exp(-\mathrm{i}T)\,(\omega-\omega_s+\omega_p)\,\exp(-\mathrm{i}\delta)\cdot\exp(-\mathrm{Q}T/2)\cdot A_+A_-[\exp(A_+\tau)-\exp(A_-\tau)]^2\}\quad(8)$$

where δ is the phase difference between the two rf fields and $A_\pm$ is defined in eq (6). To best take advantage of the interference- and uncertainty

principle-line narrowing it is convenient to measure the quantity

$$|c_s(\tau+T+\tau)_{\delta=0}|^2-|c_s(\tau+T+\tau)_{\delta=\pi}|^2.$$

We used two SPDT-solid state switches to insert into one transmission line a variable phase delay. Three counters were used to compare the count rates with both rf fields off, with both rf fields on and a relative phase difference $\delta=0$ and with $\delta=180°$. In this way the resonance in figure 5b was obtained. The exhibited line shape is typical for double oscillatory field measurements of this kind [4]. The reduction in line width is apparent.

3. Preliminary Results and Conclusion

Using one rf-spectroscopy field a number of resonance curves similar to that shown in figure 5a under various conditions and with both frequencies for the hyperfine selecting field were obtained. Since in these measurements the rf propagated along the direction of the beam, all resonances show a first order Doppler shift. The size of this correction was verified by measuring the resonances with the beam both parallel and antiparallel to the traveling rf-wave. After correcting for the Doppler shift we find for the Lamb shift in H, $n=3$ the preliminary value

$$\$(H, n=3)=315.11\pm0.89 \text{ MHz}.$$

The error is one standard deviation and is based on the statistical error and the uncertainty in the Doppler shift correction. This result is in excellent agreement with the only other measurement of this fine structure interval, which was performed by Kleinpoppen [6]. He obtained

$$\$(H, n=3)_K=313.6\pm5.7 \text{ MHz}.$$

It also agrees well with the deuterium measurement of Wilcox and Lamb [7], when the theoretical ratio $\$_H/\$_D$ is used to obtain from their deuterium value the corresponding hydrogen value. Their measurement gives

$$\$(H, n=3)_{WL}=314.93\pm0.80 \text{ MHz}.$$

All three measurements agree satisfactorily with the theoretical value [8]

$$\$(H, n=3)_{\text{theory}}=314.895\pm0.025 \text{ MHz}.$$

The success of this experiment in making fine structure measurements using a fast hydrogen beam indicates that with the double oscillatory field technique a significant reduction in the line width of the resonances can be obtained.

4. Acknowledgments

We wish in particular to acknowledge the aid of Mr. Robert Brown in constructing the accelerator and for numerous stimulating discussions. Mr. Mark Silverman contributed significantly to the reduction of the data and to the least squares fitting program.

5. References

[1] Taylor, B. N., Parker, W. H., Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[2] Robiscoe, R. T., in Cargese Lectures in Physics, edited by M. Levy (Gordon and Breach, New York, 1968), Vol. **2**, p. 3.
[3] Bashkin, S., Nucl. Instr. Methods **28**, 88 (1964).
[4] Ramsey, N. F., Molecular Beams (Oxford University Press, London, 1956), p. 115.
[5] Hughes, V. W., in Quantum Electronics, edited by C. H. Townes (Columbia University Press, New York, 1960), p. 582.
[6] Kleinpoppen, H., Z. Physik **164**, 174 (1961).
[7] Wilcox, L. R., and Lamb, W. E., Phys. Rev. **119**, 1915 (1960).
[8] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970).

DISCUSSION

H. Metcalf: Would you care to speculate on what the ultimate accuracy might be?

C. W. Fabjan: You gain a factor of 2 in linewidth for every factor of 10 you are willing to throw away in intensity. If you compare our present result and consider that we are changing our apparatus to a somewhat different configuration where we hope to get at least a factor of 100 increase in intensity, you can arrive at your own estimate.

M. Leventhal: The first question is I know you don't like magnetic fields, but would a magnetic field remove those two residual transitions?

C. W. Fabjan: You mean you could spread it out?

M. Leventhal: Yes. Get rid of it.

C. W. Fabjan: That's right. Yes.

M. Leventhal: I don't see why you should't do that. That should clean up your experiment appreciably.

C. W. Fabjan: We have heard a lot about what happens when you have velocities of the order of 10^6 centimeters per second, and it just seems to me it will be so much more difficult if you go to 10^8 centimeters per second.

M. Leventhal: Even if you try to send it down along the lines of force?

C. W. Fabjan: It will then require a very delicate alignment of your machine. On the other hand, we have some ideas why we don't get rid of the residual $F=1$ component, and we hope that in the new arrangement of our apparatus we will be able to study this in a little bit more detail.

M. Leventhal: I'd like to try another question. Bashkin and collaborators discovered when you run protons through a foil sometimes you get coherent excitation. You get coherent linear superposition of the states, which you should have in your apparatus. Does that have any relevance in your experiment at all? Can you tell me what that should do? It seems to me you ought to have sometimes linear superposition of S and P states running down your beam.

C. W. Fabjan: That's right.

M. Leventhal: You ought to have a sort of sinusoidal oscillating decay.

C. W. Fabjan: You're always far away so that the P contribution will have decayed for all practical purposes. You only see the sinusoidal effect if you apply external mixing fields.

M. Leventhal: I don't think that's true. That's wrong. That's an incorrect statement.

C. W. Fabjan: Could you say why it's wrong?

M. Leventhal: Is Bashkin here? Sellin? (*No response*)

C. W. Fabjan: Could you describe in what case you see an oscillatory behavior?

G. W. Series: I wonder if I could contribute here though I'm not in that group. I believe you would see that sort of behavior in this particular situation if you had an electric field to mix the S and the P states. You see an oscillatory behavior without any fields at all in the case of helium atoms excited to triplet P states. Then you see the triplet P, the smaller interval—whether it's 0 to 1 or 1 to 2 I can't remember. You don't need a field for that. But that's not the case with S to P transitions. The helium oscillations are all within the same multiplet. The states are all of the same parity.

M. Leventhal: Then you think he needs an external perturbation to see this alternating behavior?

G. W. Series: I do.

M. Leventhal: I apologize. Excuse me.

J. O. Stoner: Can you tell me how thick your foils are?

C. W. Fabjan: As you probably know, we get it all from the same source, and in our case it's claimed to be nominally 10 micrograms per square centimeter.

J. O. Stoner: We don't get them all from the same source. But you should keep in mind that your foil thickness will change during the course of your experiment and that as a result your velocity distribution will change. We have found at the University of Arizona that it causes some difficulty with precision measurements in this area.

C. W. Fabjan: We are aware of this. This is another reason why we wish to switch away from the foil to a gas target, to eliminate possible problems with aging of foils, etc. Also you get a much cleaner velocity distribution if you don't have to use the foil. The spread you induce by using a gas target is very small compared to the spread you get using a foil.

N. F. Ramsey: In the last slide you showed, it looked to me as if there was some asymmetry in the resonance shown, in which case, if it were there, there would be a distortion of the resonance from the peak of the curve. Do you think that asymmetry is real?

C. W. Fabjan: It's a little bit hard to decide—but it looks to me too that there is something like an asymmetry, and I think it might probably be true that we still have a little bit of the $F=1$ component in the beam.

G. W. Series: If there is trouble with overlapping components, it might be an advantage to simplify the line shape, and that might be done by somehow shaping the entry into the rf field region so that the rf field didn't come on suddenly but came on more gradually—in some way you could calculate in order to take off these wobbles on the wing to apodise the lines.

C. W. Fabjan: Are you talking now about the line shape we get with just one radiofrequency field?

G. W. Series: No, with two.

C. W. Fabjan: I think that you get this particular line shape because you extract essentially an interference phenomenon. I'm not quite sure shaping of the rf cavities will reduce this.

N. F. Ramsey: Let me comment on this. Of course, as soon as you start shaping it, then you are departing from two exactly discrete, separated oscillatory fields. And indeed I think you can do something to get rid of the wiggle. On the other hand, there are in many cases advantages to having the wiggles in that you will probably get a narrower line by virtue of allowing the wiggles to be there for a given length of the apparatus. You don't get increased narrowness when you do the shaping and have a given distance available.

G. W. Series: You might be able so to change the shape as to suppress the wing of neighboring lines, convert the Lorentzian to a Gaussian, that sort of thing.

N. F. Ramsey: That's correct, but I think usually in the process of doing that for a given total length of the apparatus the curve at the same time gets broader.

Measurement of the Lamb Shift in the $n=4$ State of Hydrogen*

R. A. Brown and F. M. Pipkin

Lyman Laboratory, Harvard University, Cambridge, Mass. 02138

Progress on a study of the fine structure in the $n=4$ state of atomic hydrogen is reported. Measurements are made as a function of the external magnetic field with a fixed radiofrequency field. A pulsed electron beam is used to excite the atoms and the signals are observed in a selected time interval following the excitation pulse. Observations are reported on transitions in both the $4S_{1/2}\rightarrow 4P_{1/2}$ and the $4S_{1/2}\rightarrow 4P_{3/2}$ complexes.

Key words: Hydrogen fine structure; Lamb shift; time resolved spectroscopy.

1. Introduction

This paper describes the progress of a study of the fine structure in the fourth excited state of atomic hydrogen. Incorporating some recently developed techniques in time-resolved spectroscopy [1], we have observed transitions between $4S_{1/2}$ and 4P sublevels in both the $j=\frac{1}{2}$ and $j=\frac{3}{2}$ manifolds. Early work [2] on the fine structure in this level is unsatisfactory because of both weak signals and a poorly-understood electric environment. The additional degree of freedom in the time domain is used to greatly enhance the signal and to explore certain systematic effects.

Figure 1, a schematic of $n=4$ in an applied magnetic field, serves to nomenclate the energy levels under investigation. A transverse rf electric field of fixed frequency is applied to this system and the magnetic field is used to adjust the separation between sublevels so that the resonant condition, $h\nu=\Delta E$, is satisfied. The S and P states have different decay probabilities and the signal is observed as a modulation of the resonance radiation originating in $n=4$. Transitions are allowed between the pairs (α, a), (β, b), (α, c), (β, d), (β, e), and (α, f), with the additional selection rule $\Delta m_I=0$.

Figure 2 shows the (α, f) and (β, e) transition frequencies with hyperfine structure included. Superimposed are other allowed transitions in $n=4$ and the cyclotron frequency for electrons, both of which seriously distort the Lamb resonances. We have tentatively chosen 250 MHz as a relatively clean frequency for studying the (α, f) transitions.

Figure 3 shows the allowed transition frequencies between the $4S_{1/2}$ and $4P_{3/2}$ sublevels. The proximity of the levels β and e and their consequently enhanced Stark coupling makes the (β, b) and (β, d) resonances unsuitable for study. We have measured the (α, c) resonances over the range of magnetic fields from 45 to 150 G.

2. Apparatus

Figure 4 is a cutaway view of the region in which the spectroscopy is performed. A 20–40 eV electron beam parallel to the magnetic field is supplied by an electron gun with a low temperature oxide cathode. It is mounted behind an optically dense, low pumping impedance baffle. The chamber is filled with about a millitorr of molecular hydrogen which the electron beam dissociates and excites, in part, to the $n=4$ level. The electron gun may be operated in a fast-pulsed or d-c mode, and the current is regulated with a long time constant.

The photomultiplier tube, RCA 8575, monitors the visible photons associated with spontaneous decays from the $n=4$ to the $n=2$ level, H_β. A lens collimates the light, which then passes through an interference filter. Early work was done with a 100 Å wide filter; a new 10 Å filter has improved the signal by a factor of 5, presumably due to the exclusion of molecular radiation. Pulse height discrimination is performed on the photomultiplier output and the single photon events are processed in 100 MHz coincidence and counting circuitry.

Figure 5 shows the full vacuum system and its position inside the Helmholtz pair which provides the magnetic field. The voltage drop across a standard resistance in series with the solenoid windings is monitored with a digital voltmeter to measure the field. Calibration was performed independently using a Series 820 Rawson-Lush probe and the cyclotron resonance; it is believed accurate to better than 0.1 percent.

The vacuum system is entirely stainless steel and bakeable to 450 °C. An ion pump is employed to avoid oil films which can be the source of troublesome electric perturbations and cause difficulties with the

* Research supported by NSF Grants GP-8551 and GP-22787.

electron gun. Between runs the system is thoroughly outgassed: with the gun on, the ultimate pressure is below 10^{-8} torr. Hydrogen gas is bled into the chamber through the variable leak, while the pressure is measured using a Varian MilliTorr ionization gage. The strongest signals are observed with hydrogen pressures between 0.5 and 2.0 millitorr (mtorr).

General Radio unit oscillators are used to supply the radio frequency field. Frequencies below 500 MHz ($S_{1/2}$–$P_{1/2}$ transitions) are measured directly on a frequency counter with an internal converter.

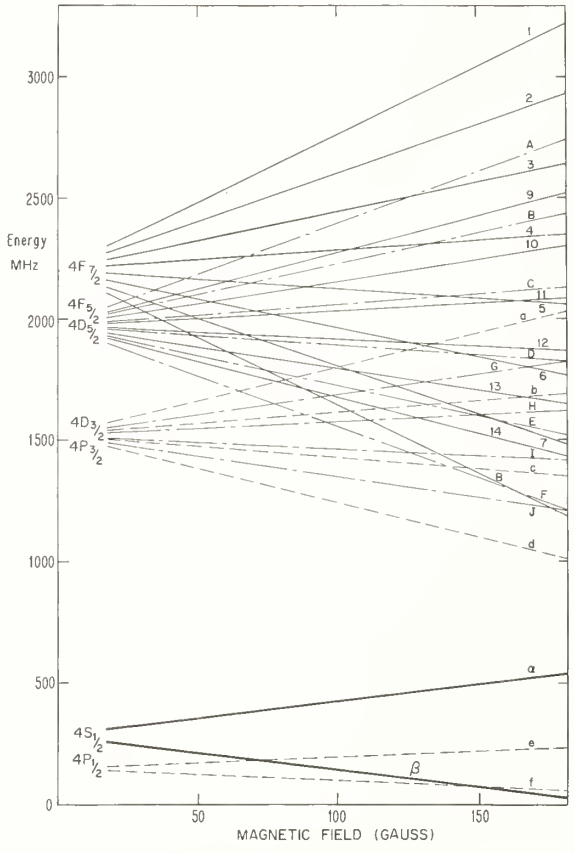


FIGURE 1. *The energy levels in the fourth excited state of atomic hydrogen, neglecting hyperfine structure and* L–S *decoupling.*

Those frequencies around 1 GHz ($S_{1/2}$–$P_{3/2}$ transitions) are mixed with a 1.45 GHz standard frequency and the beat is measured. The rf electrodes are a capacitive load inserted in a transmission line, which is terminated in a sampling scope for voltage measurements.

3. Experimental Design

Figure 6 shows one of the three logical schemes in which data have already been taken. An electron pulse about 300 ns long populates the $n=4$ level.

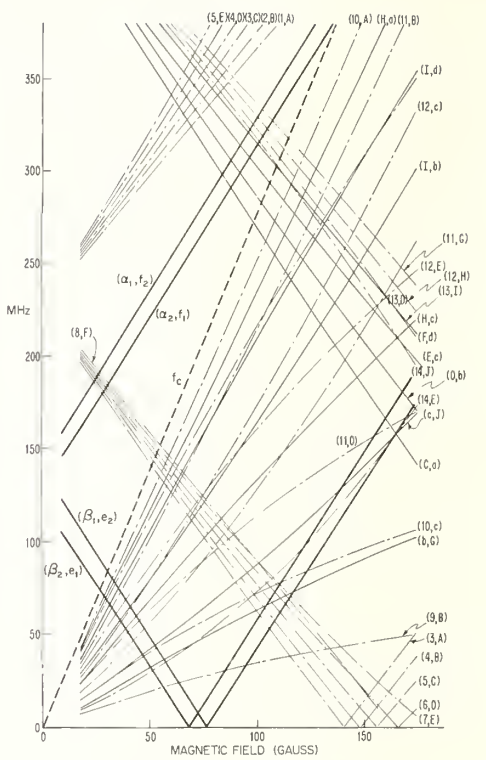


FIGURE 2. *The* $4S_{1/2}-4P_{1/2}$ *transition frequencies in a transverse rf electric field.*

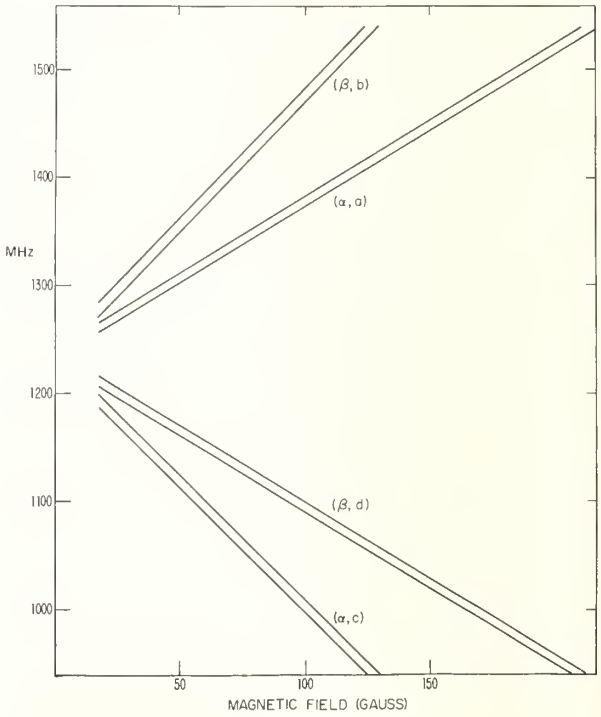


FIGURE 3. *The* $4S_{1/2}-4P_{3/2}$ *transition frequencies in a transverse rf electric field.*

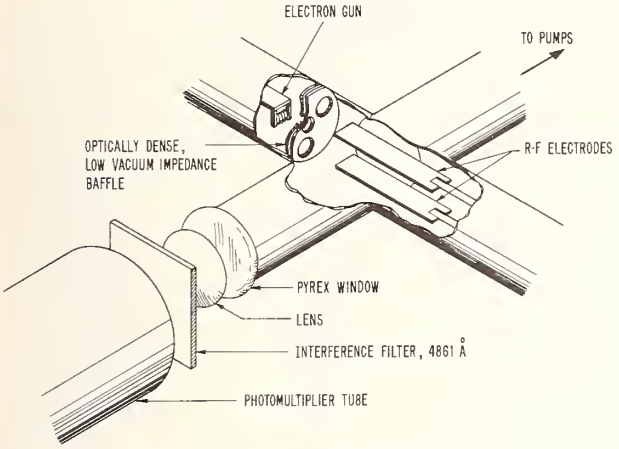


FIGURE 4. *A cutaway view of the portion of the vacuum system in which the spectroscopy is performed.*
Resonance radiation from electron-impact excited hydrogen atoms is monitored by the photomultiplier tube.

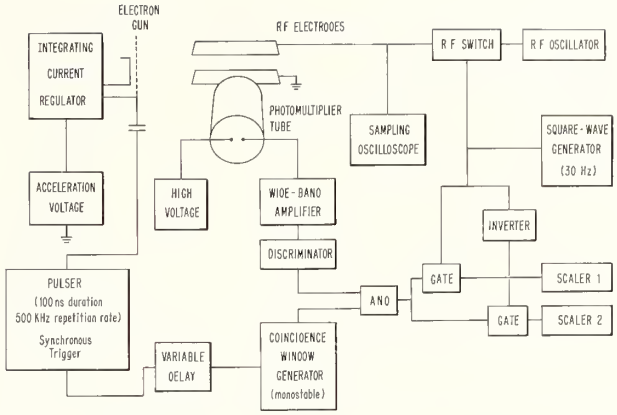


FIGURE 6. *A block diagram of the primary experimental design.*
The coincidence network selectively monitors the 4S-state population.

Each of the four angular momentum states dies away at its own rate, but the S-states are by far the longest lived, with a mean life of 230 ns. The lifetimes of the P, D, and F states are 12, 36, and 73 ns respectively, and the F states cannot contribute to the H_β radiation. A coincidence network selects only those photomultiplier events which occur in time out on the tail of the decay, which is rich in S-states and may be free from any perturbing effect of the electron beam. The coincidence window is typically 300 ns long and delayed by about 100 ns from the end of the excitation pulse. The repetition rate for the whole cycle is 500 kHz. A low frequency square wave modulates the rf field and also alternates the sealer in which the photomultiplier events are recorded. This isolates the effect due to spectroscopic transitions as a difference in the count rates of the two scalers. The signal is taken as the fractional difference in the two scaler readings. We run at one magnetic field until the desired statistics are obtained.

A second arrangement pulses on the rf in time only after the eleetron excitation pulse has passed. A third uses a continuous electron beam, bypasses the coincidence network and uses cw rf just as in the first setup.

The signal using a pulsed beam is an order of magnitude higher than with a d-c beam. Completely saturated resonances as strong as 20 percent have been observed.

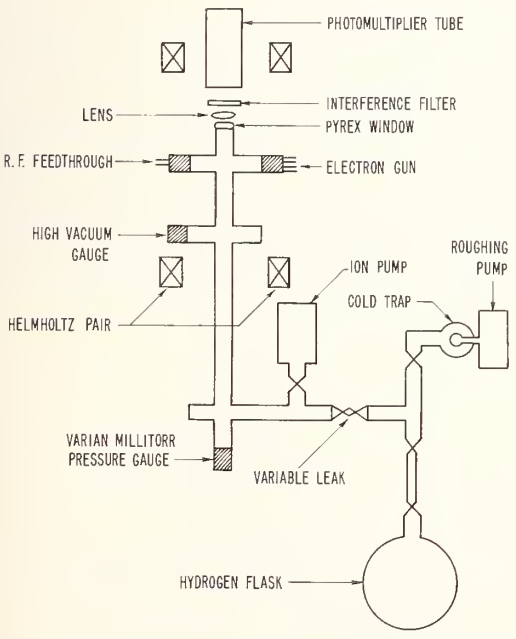


FIGURE 5. *A schematic diagram of the vacuum system, showing its position inside the magnet.*

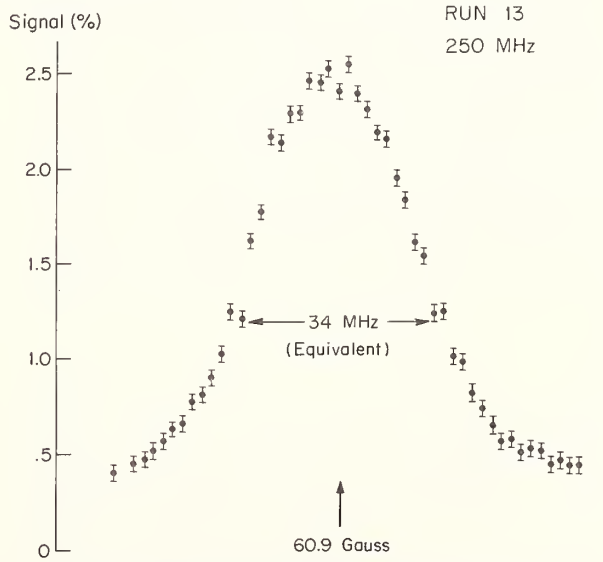


FIGURE 7. *A typical* (α, f) *resonance curve.*
There are two unresolved resonances about 15 MHz apart.

TABLE 1. *A summary of the* (α, f) *measurements at 250 MHz*

Run	H press (mtorr)	Field orientation	$V_{RF_{PP}}$ (volts)	Beam mode	$V_{Acc.}$ (volts)	RF mode	Analysis method	Line center (gauss)
1	0.8	+	1.4	DC	25	CW	C splitt	60.88±0.26
2	0.8	+	1.3	DC	25	CW	C splitt	60.96±0.26
3	0.8	+	1.3	PLSD	25	CW	C splitt	61.17±0.12
4	0.8	+	1.0	PLSD	25	CW	Symm pts	61.23±0.26
5	0.8	−	1.5	PLSD	25	CW	C splitt	61.30±0.09
6	0.8	−	1.6	PLSD	25	CW	Symm pts	61.02±0.09
7	0.8	−	1.0	PLSD	25	CW	Symm pts	60.81±0.11
8	0.8	−	1.5	PLSD	25	PLSD	Symm pts	61.27±0.16
9	0.6	−	0.8	PLSD	25	CW	Symm pts	60.61±0.18
10	5.0	−	1.0	PLSD	25	CW	C splitt	60.95±0.18
11	0.5	−	1.0	PLSD	25	CW	C splitt	60.00±0.18
12	0.5	+	1.0	PLSD	35	CW	C splitt	60.98±0.18
13	0.5	+	1.0	PLSD	35	CW	C splitt	60.88±0.18
14	0.8	+	—	PLSD	35	CW	Symm pts	61.01±0.09

4. The $(\alpha$-$f)$ Resonances

Figure 7 shows a typical (α, f) resonance. Due to hyperfine structure there are two unresolved peaks, each about 13 MHz wide and separated by 15 MHz. Figure 8 shows a typical dependence of the signal at resonance on rf voltage as measured at the sampling scope. Measurements are normally made with the resonance less than 50 percent saturated.

One approach to finding the midpoint of the resonance is to split a panoramic curve such as this at several heights, averaging the results. Another approach is to confine the measurements to nearly symmetrical points and to use the slope to correct one's guess. Both techniques have been employed in the (α, f) measurements and the results are compatible. Once it is established that the resonances are undistorted and that the resonance centers can be reliably predicted with 1 percent accuracy, the latter technique emerges as the most efficient approach. All the (α, c) measurements have been taken by symmetrical points.

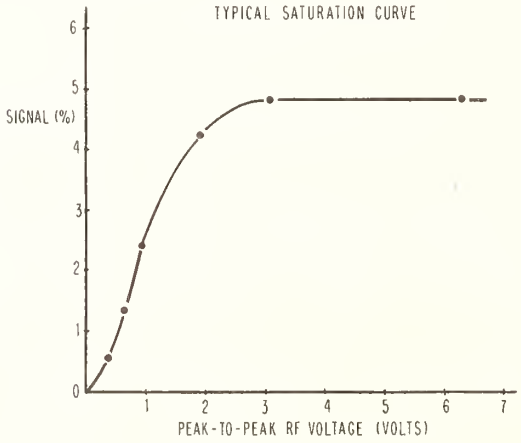


FIGURE 8. *A typical saturation curve.*
Measurements were normally taken with the resonance less than 50 percent saturated.

Table 1 is a summary of the work to date on the (α, f) transition at 250 MHz. The motivation at this stage was to expose any experimental parameter which affects the position of the resonance and its height and width. The gas pressure may play a role in determining the local electric field, and thereby the level positions through the Stark effect. Data were taken and field calibration was performed with the magnetic field parallel and antiparallel to the electron beam. The rf power not only affects the shape of the Lamb resonance, but also affects the nearby cyclotron resonance. The electron energy may affect the average velocity of the atomic dissociation products and thereby the average motional electric field [3].

Figure 9 is a histogram of these measurements. The first observation is that the results are consistent. On the other hand there is evidence of a strong systematic effect: the resonance is shifted by more than 1 MHz from the position predicted by the theoretical values for the zero-field splittings [4] and the known magnetic field dependence [5]. The stability of the resonance under this wide variety of experimental conditions implies that the effect is not associated with one of the controllable experimental parameters and may have its origin in the motional electric field seen by the atoms. Such a field

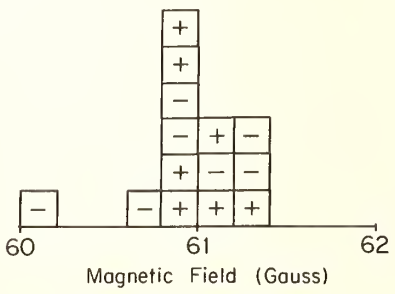


FIGURE 9. *A histogram of the* (α, f) *measurements at 250 MHz summarized in table 1.*
The sign indicates the orientation of the magnetic field.

Table 2. *A summary of the* (α, c) *measurements at various magnetic fields*

Frequency (MHz)	Number of measurements	Resonance center (gauss)
1135.1	4	43.91±0.06
1124.4	3	48.52±0.04
1110.4	3	54.41±0.04
1110.2	2	54.56±0.04
1109.8	5	54.80±0.05
1088.6	5	64.08±0.09
1066.1	4	73.60±0.09
1040.5	5	84.74±0.03
1015.9	4	95.39±0.09
997.3	8	103.47±0.13
968.1	3	116.28±0.04
930.0	4	132.99±0.07
900.1	5	146.38±0.14

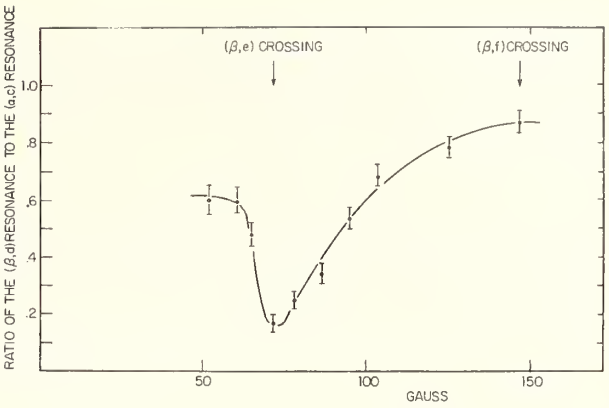


Figure 11. *The ratio of the height of the* (β, d) *resonance to that of the* (α, f) *resonance as a function of magnetic field.*

is strictly transverse and needs to be of the order of 1 V/cm. The implied transverse component of atomic velocity is about 2×10^6 cm/s. Because undistorted (α, f) resonances are obtainable only over a narrow magnetic field range, it is left to the (α, c) resonances to explore this effect.

5. The (α-c) Resonances

A single scheme has been used in the measurement of the (α, c) transitions. A pulsed beam with cw rf is used as illustrated in figure 6. The operating pressures are between 0.7 and 0.9 mtorr, and the acceleration voltage is 36 eV. One measurement consists of eight pairs of observations of the heights of closely symmetrical points on the two sides of the resonance, which are averaged and corrected with the slope.

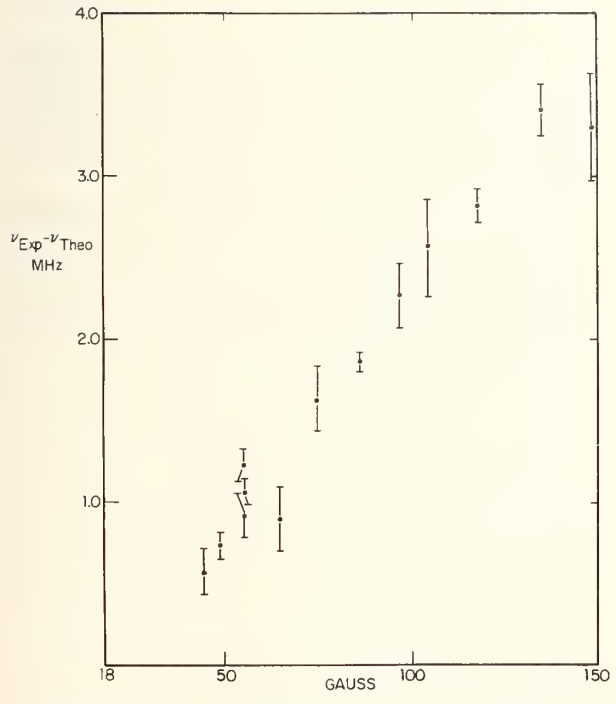


Figure 10. *The difference between the theoretical and measured* (α, c) *transition frequencies as a function of magnetic field.*

Table 2 is a summary of the (α, c) measurements. Their deviation from the positions predicted by the theoretical values of ΔE and S and the Zeeman effect shows a marked dependence on the magnetic field, illustrated in figure 10. Theoretical estimates of the resonance center are determined in the following way: transition frequencies, computed by diagonalizing the Hamiltonian of Brodsky and Parsons [5], are combined in a Lorentzian line shape of the appropriate full height and width. Points of a given height on the two sides of the resonance are found by an iterative procedure and averaged to find the center.

A strong indication that a transverse electric perturbation is playing a role here is supplied by comparing the strengths of the resonance (α, c) and (β, d). The level β crosses the level e at about 70 G and the level f at about 150 G. At the first crossing a transverse electric field will deplete the β population, while a longitudinal field must serve at the second crossing. Figure 11 shows that the (β, d) resonance, which reflects the initial β population, is profoundly affected at the first crossing, while slightly if at all, at the second crossing. A transverse electric field of the order 1 V/cm is required to completely saturate the (β, e) resonance.

6. Summary

A procedure has been established to measure strong and stable fine structure resonances in the fourth excited state of atomic hydrogen. Relatively undistorted line shapes have been obtained over a range of magnetic fields for the $S_{1/2}$–$P_{3/2}$ transitions, while only at one center field value for the $S_{1/2}$–$P_{1/2}$ transitions. A systematic shift of the resonance position of the order 1 percent of the Lamb shift is found, which is dependent on the magnetic field.

Following the implication that the motional electric field may be the origin of this effect, analysis is proceeding that will seek to fit all the observations with a single velocity distribution.

7. Acknowledgment

The authors wish to express their appreciation to Claude Bernard for his assistance in the measurements and in the preparation of this paper.

8. References

[1] Hatfield, L. L., and Hughes, R. H., Phys. Rev. **156**, 102 (1967).
[2] Lamb, W. E., Jr., and Sanders, T. M., Jr., Phys. Rev. **119**, 1901 (1960). Wilcox, L. R., and Lamb, W. E., Jr., Phys. Rev. **119**, 1915 (1960).
[3] Leventhal, M., Robiscoe, R. T., and Lea, K. R., Phys. Rev. **158**, 49 (1967).
[4] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[5] Brodsky, S. J., and Parsons, R. G., Phys. Rev. **163**, 134 (1967).

DISCUSSION

M. Leventhal: I was so unlucky last time that I want to try again. Could I have your last slide for a second?

R. A. Brown: Certainly. Could we have the last slide?

M. Leventhal: On the last slide you quote a value for the transverse velocity distribution. Is that not the velocity of the hydrogen atom produced in dissociation of molecular hydrogen by your electron bombardment?

R. A. Brown: Is it not?

M. Leventhal: Isn't that how you're doing it?

R. A. Brown: The dissociative process is assumed to be the origin of the velocity.

M. Leventhal: How do you get such a high velocity?

R. A. Brown: That is the velocity assumed in this fit.

M. Leventhal: Yes, but we have made a measurement of the velocity of hydrogen metastable atoms made in exactly that fashion.

R. A. Brown: That's right, but that was done in the $n=2$ state, not in the $n=4$ state.

M. Leventhal: I know, but it shouldn't be that different.

R. A. Brown: Shouldn't it? I don't know.

M. Leventhal: I don't think it should be, but I've been wrong before. (*Laughter*) Oh, that value just gives you the best fit. Is that what you're saying?

R. A. Brown: Well, it was a by-hand calculation, and it's not a complete least squares fit by any means. It's just a calculation of those two points where the X's are, those two values, and, of course, it has to go through the origin.

M. Leventhal: From what I know of the problem, I think it ought to be probably three or four times smaller than that.

R. A. Brown: You're saying there should be no difference when you go to $n=4$ at all?

M. Leventhal: There might be small differences but not like that I don't think. One other question I would like to ask is how you normalized your data as you tune magnetic fields to the resonance. I didn't pick it up.

R. A. Brown: Well, the lines are symmetric far beyond the uncertainty. In other words, there's no noticeable asymmetry in these lines over the range in which we use them, so we are just determining the line centers.

M. Leventhal: I see. So you just assume that normalization wasn't necessary basically?

R. A. Brown: I'm not sure what you mean by "normalization."

M. Leventhal: I mean when you tune to a resonance as you vary your magnetic field, people that have done similar experiments find that, for instance, the number of atoms you make in the state of interest changes as a function of magnetic field, and this—

R. A. Brown: The number of atoms?

M. Leventhal: In the relevant state. For various mysterious reasons.

R. A. Brown: Well, then this should show up as an asymmetry in the line.

M. Leventhal: Yes, it should. I'm surprised that you don't see any asymmetry at all.

Determination of the Fine Structure Constant α From Helium Fine Structure*

A. Kponou, V. W. Hughes, C. E. Johnson,** S. A. Lewis,*** and F. M. J. Pichanick****

Physics Department, Yale University, New Haven, Conn. 06520

A measurement of ν_{01}, the interval between the components $J=0$ and $J=1$ of the fine structure in the 2^3P state of helium, has been made by an optical-microwave atomic beam magnetic resonance technique. The result is $\nu_{01}(\text{expt.})=29{,}616.864\pm0.036$ MHz (±1.2 ppm). The (one standard deviation) error is determined primarily by the statistical counting error. This result is 3.09 MHz smaller than the present theoretical value. This difference is presumably due to an uncalculated theoretical term of $O(\alpha^4 Ry)$. When this latter calculation is completed it should be possible to determine α to 1 ppm.

Key words: Fine structure constant; helium fine structure.

1. Introduction

An important possibility for determining the fine structure constant α to a precision of about 1 part per million (ppm) is the study of the fine structure intervals in the 2^3P state of helium [1, 2]. The advent of high-speed digital computers has made possible the calculation of accurate nonrelativistic wavefunctions for helium and hence also of accurate values of the fine structure intervals. Precise measurements of the intervals have been made by the optical-microwave atomic beam magnetic resonance method. The fine structure constant α can be determined by equating the theoretical expression for a fine structure interval, in which α is treated as unknown, with the experimental value for the same interval. The uncertainty of the determination depends, of course, on the uncertainties of both the measured values and the calculated terms.

The 2^3P state of helium was chosen for this study for several reasons. Since the lifetime of this state is 10^{-7} s as compared to the value of 1.6×10^{-9} s for the 2^2P state of hydrogen, and also since the larger fine structure interval $J=0$ to $J=1$ of helium is nearly three times larger than the fine structure interval $2^2P_{3/2}$–$2^2P_{1/2}$ of hydrogen, the ratio of interval to natural line width is about 180 times larger for helium than for hydrogen. Hence a more precise measurement of the helium fine structure interval can be expected, and indeed has been achieved. The 2^3P state is the lowest state of a particular symmetry (spin and orbital angular momentum), and hence variational calculations can be most easily applied to this state. The disadvantage of the use of helium as compared to hydrogen for the determination of α is the relative complexity of the theory of the two-electron atom.

2. Theory of Fine Structure

Recently, the modern theory of helium fine structure has been reviewed [2]. The difference between the energy E_J of the fine structure level with total angular momentum J and the Schroedinger energy E can be written:

$$E_J-E=\langle O\mid \mathcal{H}_2\mid O\rangle_J+O(\alpha^3)+\langle O\mid \mathcal{H}_2(\mathcal{H}_0-E)^{-1}\mathcal{H}_2\mid O\rangle_J+\langle O\mid \mathcal{H}_4\mid O\rangle_J+\cdots;\qquad J=0,1,2. \quad (1)$$

The Schroedinger equation is $\mathcal{H}_0\mid O\rangle=E\mid O\rangle$. The leading term is the expectation value of the Breit interaction operator $\alpha^2\mathcal{H}_2$ which gives the spin-orbit and spin-spin interactions and is of the order $\alpha^2 Ry$. The second term is the lowest order virtual radiative correction, due to the electron anomalous magnetic moment [3, 4]. The third term is the contribution of $\mathcal{H}_2$ in second-order perturbation and is of order $\alpha^4 Ry$. The fourth term includes the remaining relativistic and virtual radiative contributions of order $\alpha^4 Ry$. The leading term of order $\alpha^2 Ry$ has been computed to an accuracy of 1 to 2 ppm [5, 6], and the order $\alpha^3 Ry$ term has been evaluated [3, 4]. Calculations have been done and are still in progress on the two $\alpha^4 Ry$ terms [7, 8].

* This research supported by the Air Force Office of Scientific Research, United States Air Force under AFOSR Grant No. F44620-71-C-0042.

** Present address: Physics Department, University of California, Berkeley 94720.

*** Present address: Physics Department, University of California, Davis 95616.

**** Present address: Physics Department, University of Massachusetts 02167.

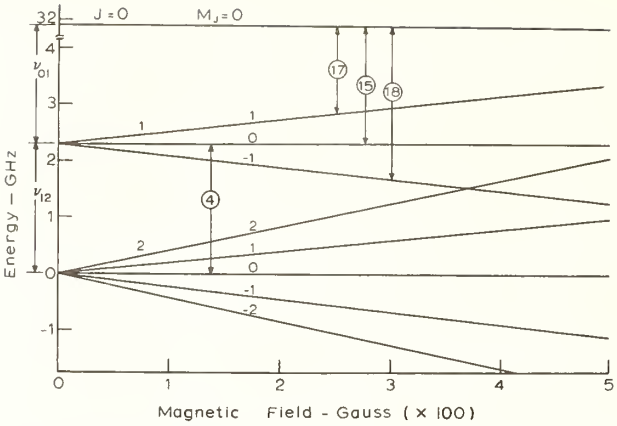


FIGURE 1. *Low-field Zeeman effect in the* 2^3P *state of* He^4.
Note break in energy axis. Transitions "4" and "15" studied to measure ν_{12} and ν_{01} respectively.

3. Atomic Beam Experiment

The fine structure intervals have been measured using the optical-microwave atomic beam magnetic resonance technique [9]. Figure 1 shows the Zeeman energy level diagram of the 2^3P state of helium. The transitions which have been studied to determine ν_{01} and ν_{12} are 15 at a field of 300 G, and 4 at 100 G and 500 G. Since these transitions are both $m_J = 0 \rightarrow m_J' = 0$ at weak magnetic field, the transition frequencies are independent of magnetic field to first order.

The scheme of the experiment is shown in figure 2. A beam of 1^1S_0 ground state helium atoms is formed when atoms effuse into the vacuum chamber through a narrow slit. An electron gun next to the source cell excites these atoms to the metastable 2^3S state by electron bombardment with an efficiency of about 1/5000. The inhomogeneous A- and B-magnetic fields, together with the collimator and beam stops, select atoms with a particular m_J component of the $^3S_1$ state. In the C-region are the homogeneous static magnetic field H_0, a helium lamp to induce the optical transition $2^3S \rightarrow 2^3P$, and a microwave field to induce the transition 4 or 15. The apparatus is

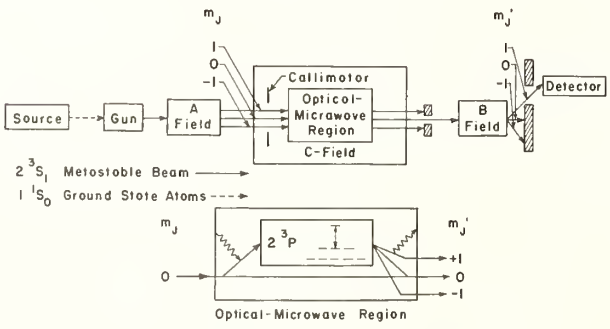


FIGURE 2. *Scheme of experiment with stops and detector arranged for* $2^3S_{m_J=0\rightarrow1}$ *detection (top).*
The three processes in the optical microwave region are optical excitation and decay (wiggly arrows), and rf mixing (double arrow).

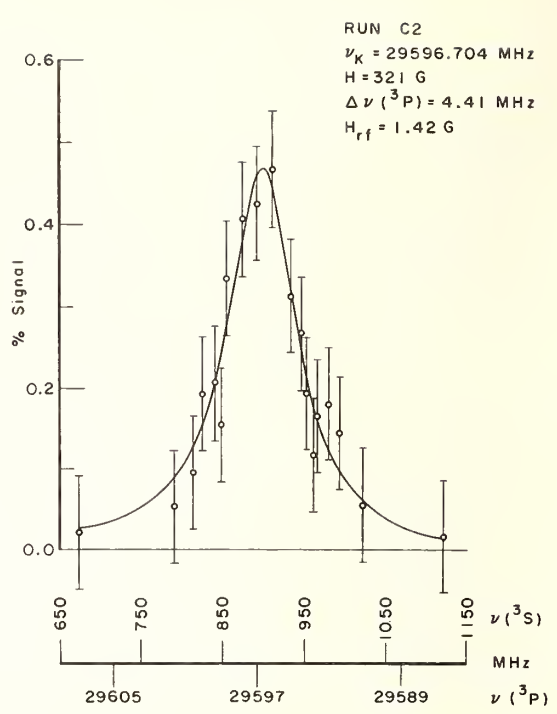


FIGURE 3. *Observed resonance and Lorentzian fit.*
Line width $\Delta\nu(^3P)$ is approximately 44 G. $\overline{H}_{rf}$ is an average of the sinusoidal rf field in the interaction region. ν_K is the microwave klystron frequency.

arranged to observe transition 4 or 15 by detecting atoms initially in $m_J = 0$ of the 2^3S_1 state in the A-region and finally in $m_J = \pm 1$ in the B-region. The signal is taken as the (normalized) difference between the total counts at the detector with the microwave power on and off. Resonance curves are obtained by varying the magnetic field with the microwave frequency fixed.

TABLE 1. 2^3P *fine structure of helium.*

Experiment Method	ν_{12}(MHz)	ν_{01}(MHz)
Polarization of decay radiation[a]	2291.72±0.36	
Level-crossing[b]	2291.200±0.022	29616.76±0.40
Atomic beam	2291.196±0.005[c]	29616.864±0.036[d]
Theory		
Schwartz[e]	2289.618	29619.910
Pekeris et al.[f]	2289.645	29619.933

[a] I. Weider and W. E. Lamb, Jr., Phys. Rev. **107**, 125 (1957).
[b] J. R. Lifsitz, Ph.D. Thesis, University of Michigan (1965).
[c] F. M. J. Pichanick et al., Phys. Rev. **169**, 55 (1968).
[d] This experiment (to be published).
[e] C. Schwartz, Phys. Rev. **134**, A1181 (1964). The value $\alpha^{-1} = 137.03602$ (21) was used.
[f] B. Schiff, C. L. Pekeris, and H. Lifson, Phys. Rev. **137**, A1672 (1965).

More detailed descriptions of the experimental method and apparatus have been given [9, 10]. The determination of ν_{12} has been completely discussed [9].

A typical resonance line is shown in figure 3 for the determination of ν_{01}. The line width of approximately 4.4 MHz is due to the radiative natural line width and to microwave power broadening. The curve is a Lorentzian fit to the observed points. Seventy-three resonance curves were obtained.

In the analysis of the data the following systematic errors were considered:

(1) Static magnetic field homogeneity;
(2) Microwave frequency and power stability;
(3) Variation of microwave magnetic field over transition region;
(4) Detector stability;
(5) Background counting effects;
(6) Effect of variation of H_0 on optical and microwave transition probabilities;
(7) Doppler effect;
(8) Bloch-Siegert effect;
(9) Effect of nearby resonances.

The experimental result is $\nu_{01}=29616.864\pm0.036$ MHz (1.2 ppm). The error represents 1 standard deviation and is due mainly to counting statistics. The precision is an order of magnitude improvement over the level crossing result [11].

Table 1 lists the various theoretical and experimental values for the 2^3P fine structure intervals. The present large discrepancies between the theoretical and experimental values are presumably due to the uncalculated $\alpha^4 Ry$ terms in eq (1). When these terms are calculated, it should be possible to determine α to about 1 ppm.

4. References

[1] Hughes, V. W., Quantum Electrodynamics: Experiment, Atomic Physics, ed. V. W. Hughes, B. Bederson, V. W. Cohen and F. M. J. Pichanick, Plenum Press, New York (1969), p. 15.
[2] Schwartz, C., Precise Schroedinger Wave Functions for Two Electrons [1], p. 71.
[3] Araki, H., Prog. Theoret. Phys. (Kyoto) **17**, 619 (1957).
[4] Sucher, J., Phys. Rev. **109**, 1010 (1958).
[5] Schwartz, C., Phys. Rev. **134**, A1181 (1964).
[6] Schiff, B., Pekeris, C. L., and Lifson, H., Phys. Rev. **137**, A1672 (1965).
[7] Araki, G., Ohta, M., and Mano, K., Phys. Rev. **116**, 651 (1959).
[8] Hambro, L., thesis, University of California, Radiation Laboratory Report No. 19328 (1969), and private communication; Kim, K. Y., Phys. Rev. **140**, A1498 (1965).
[9] Pichanick, F. M. J., Swift, R. D., Johnson, C. E., and Hughes, V. W., Phys. Rev. **169**, 55 (1968).
[10] Lewis, S. A., Pichanick, F. M. J., and Hughes, V. W., Phys. Rev. **A2**, 86 (1970).
[11] Lifsitz, J., and Sands, R. H., Bull. Am. Phys. Soc. **10**, 1214 (1965).

Precision Measurement of the Lamb Shift in Singly Ionized Helium*

Musti A. Narasimham**

National Bureau of Standards Boulder Laboratories and The University of Colorado, Boulder, Colo. 80302

and

R. L. Strombotne***

National Bureau of Standards Boulder Laboratories, Boulder, Colo. 80301

The Lamb shift (S) in the $n=2$ level of singly ionized helium has been measured by using a microwave method. The experiment has been performed in an applied homogeneous magnetic field at three different operating values for the helium pressure and the result is extrapolated to a zero pressure value. A search has been made for all possible systematic errors in the experiment. The result for S thus obtained is 14046.2 ± 2.0 MHz. This result agrees with the previous experimental values (within three standard deviations) and the theoretical value of 14044.5 ± 5.2 MHz.

Key words: Fine structure; Lamb shift; singly ionized helium.

1. Introduction

Lamb and his collaborators [1] have studied the fine structure of the $n=2$ level of hydrogen and deuterium, using an atomic beam, radiofrequency power method, where they not only established the existence of a small energy separation between the $2^2S_{1/2}$ and $2^2P_{1/2}$ levels of these atoms but also measured precise values for this separation. This separation (the Lamb shift, S) has a pure radiative origin, and hence, provided a testing ground for the validity of the theory of quantum electrodynamics. A thorough compilation of all Lamb shift measurements made by several researchers, including those that utilized level-crossing techniques, is contained in Taylor et al. [2]. The best value reported for S in the $n=2$ level of singly ionized helium (prior to this work) is due to Lipworth and Novick [3], and is 14040.2 ± 4.5 MHz. The quoted uncertainty includes three standard deviations for statistics plus estimated uncertainty due to systematics.

This report discusses a measurement of the Lamb shift in singly ionized helium in the $n=2$ level, and the experiment uses rf power to induce the Lamb shift transition, with a phase-sensitive detection scheme in signal measurements.

2. Experimental Method

Helium gas contained in the space between the two reflecting mirrors of a spherical mirror Fabry-Perot resonator is bombarded continuously by electrons accelerated through 200 V. A small fraction of the atoms is ionized and excited to the $2^2S_{1/2}$ state of interest (typically, one percent of ion content). The energy level diagram relevant to the atom and ion (up to $n=4$ level) is shown in figure 1, while figure 2 shows the fine structure of the $n=2$ level of the ion. The ions in the $2^2P_{1/2}$ level are short lived, with a lifetime of 10^{-10} s. The presence of rf power of energy appropriate to the $2^2S_{1/2}$–$2^2P_{1/2}$ separation induces transitions in the metastable $2^2S_{1/2}$ ions to the $2^2P_{1/2}$ states which will then rapidly decay to the ionic ground state $1^2S_{1/2}$ with the emission of photons whose wavelength is about 304 Å.

The experiment is performed in an external magnetic field, in which $2^2S_{1/2}$, $2^2P_{1/2}$ and $1^2S_{1/2}$ levels split into the magnetic substates α, β, e, f, and α', β' respectively (fig. 3 shows the Zeeman effect of $2^2S_{1/2}$ and $2^2P_{1/2}$ levels). We have chosen to work with the $\alpha\rightarrow e$ transition at a magnetic field value close to the $\beta\rightarrow f$ level-crossing field value. Since the $\alpha\rightarrow e$ transition is a π transition, microwave power with its electric field vector parallel to the applied magnetic field direction is fed into the resonator. The α state ions transferred by induced transitions to the e state decay rapidly to the α' and β' states. The uv photon flux arising from the latter transitions forms a small addition to a large background (typically 0.3 percent).

---

* Based upon work submitted by Musti A. Narasimham in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Physics and Astrophysics, University of Colorado, Boulder, Colo. 80302.

** Present address: IBM Corporation, East Fishkill Facility, Hopewell Junction, N.Y. 12533.

*** Present address: Department of Transportation, Washington, D. C. 20590.

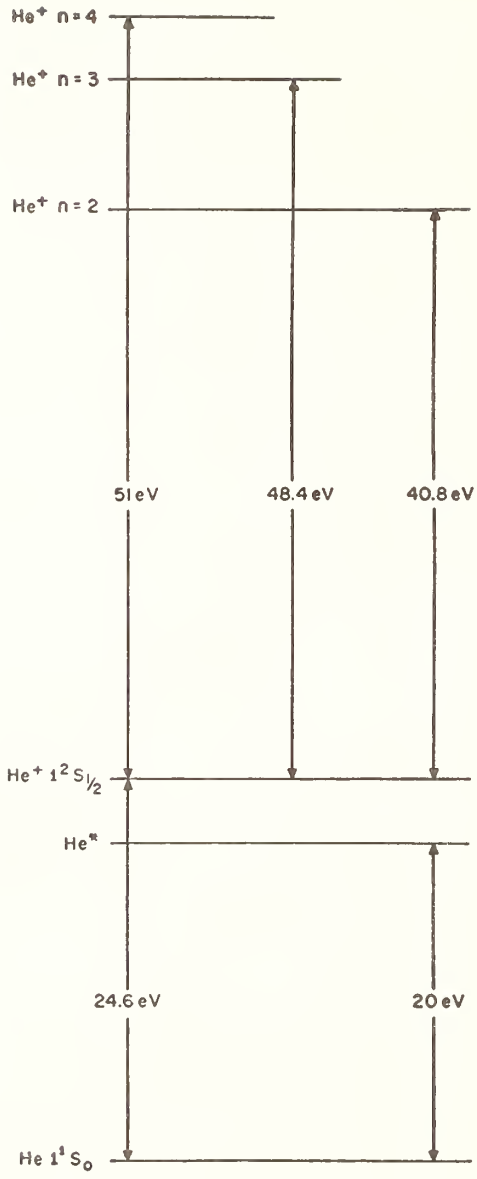


FIGURE 1. *Energy levels of* He *and* He^+ *(note that the* He^+ *levels are not complete).*

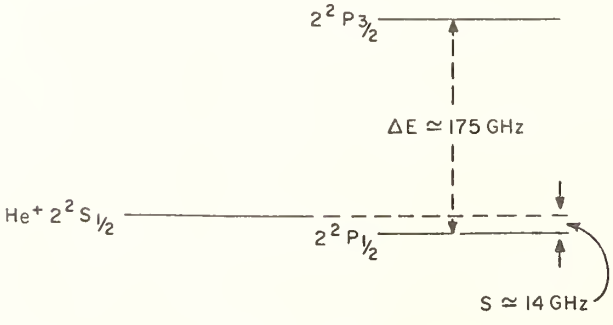


FIGURE 2. *Fine structure of singly ionized helium,* n=2 *level.*

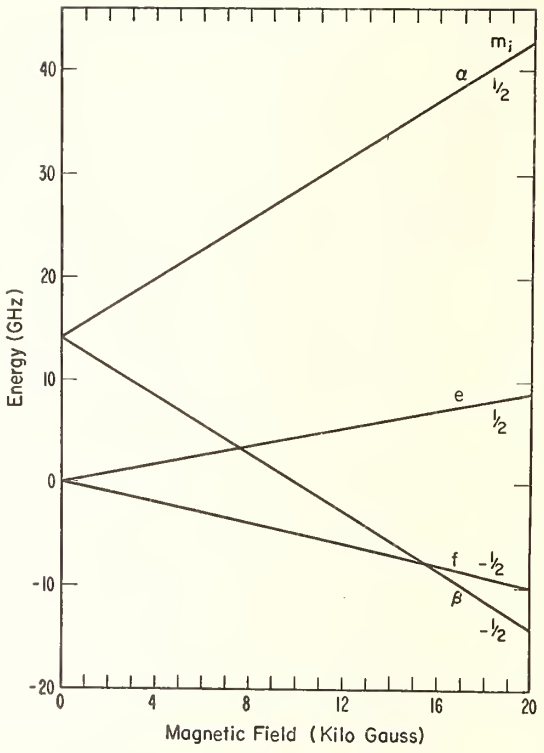


FIGURE 3. *Zeeman effect of the* $2^2S_{1/2}$ *and* $2^2P_{1/2}$ *levels of the* He^+ *ion.*

α, β, and e, f are the magnetic substates of S and P levels, respectively.

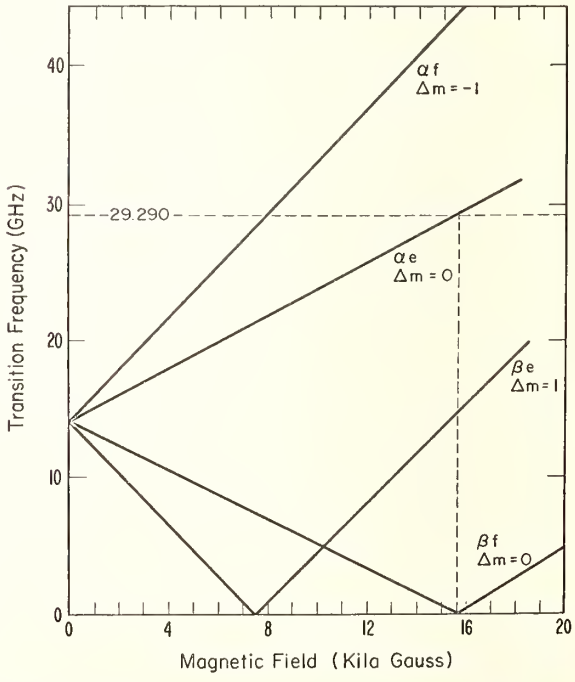


FIGURE 4. *Zeeman transition frequencies as function of magnetic field.*

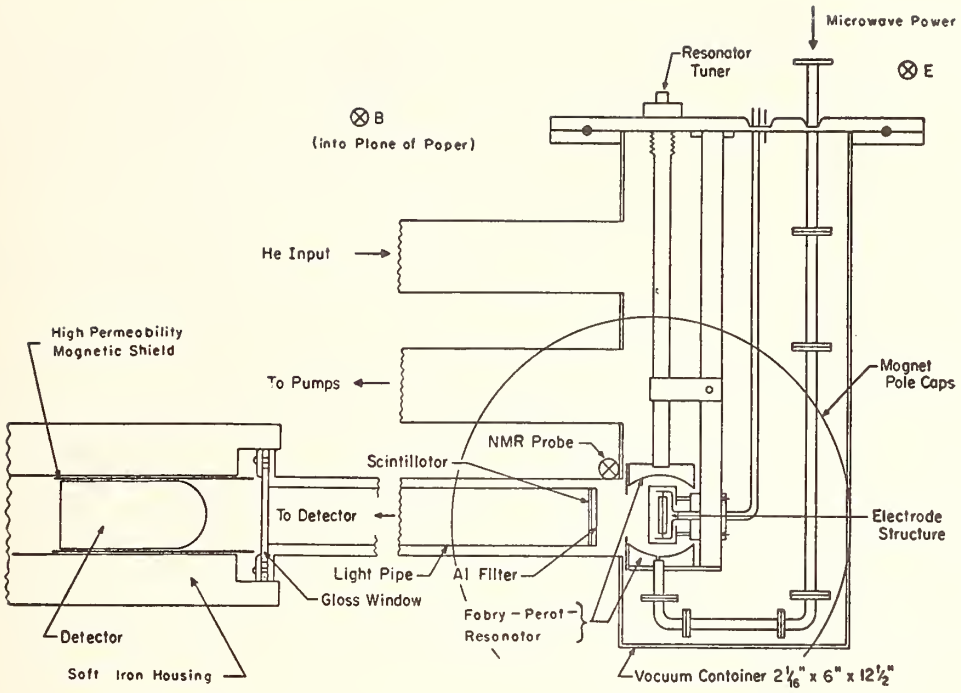


FIGURE 5. *Schematic of the main experimental set up.*

The background is due to de-excitation of short-lived excited ionic and atomic states, plus leakage visible light from the filament that provides the bombarding electron current. The uv signal photons that escape in the direction of the detector pass through an aluminum filter, strike a sodium salicylate scintillator where they are quantum-converted to visible photons (≃4000 Å wavelength), which are collected by a light pipe and guided to a photodetector. The rf power is modulated at a 50 Hz rate, and the signal is detected synchronously by a phase-

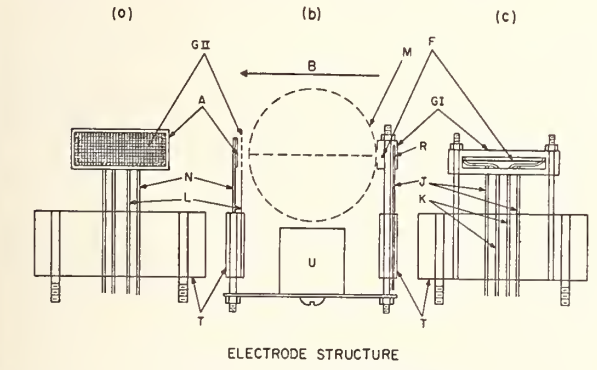


FIGURE 6. *Electrode structure.*
F is maintained 200 V negative, GI, GII, and R are held at ground and *A* is held +45 V.

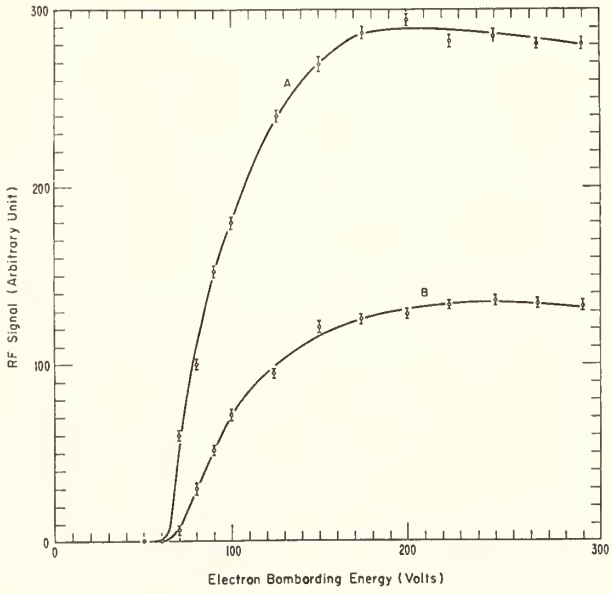


FIGURE 7. *Excitation curves for the rf induced signal:* (A) *4.6 mtorr* (B) *2.3 mtorr for the pressure.*
The sharp rise in signal at about 65 V is due to the onset of excitation of the metastable α states.

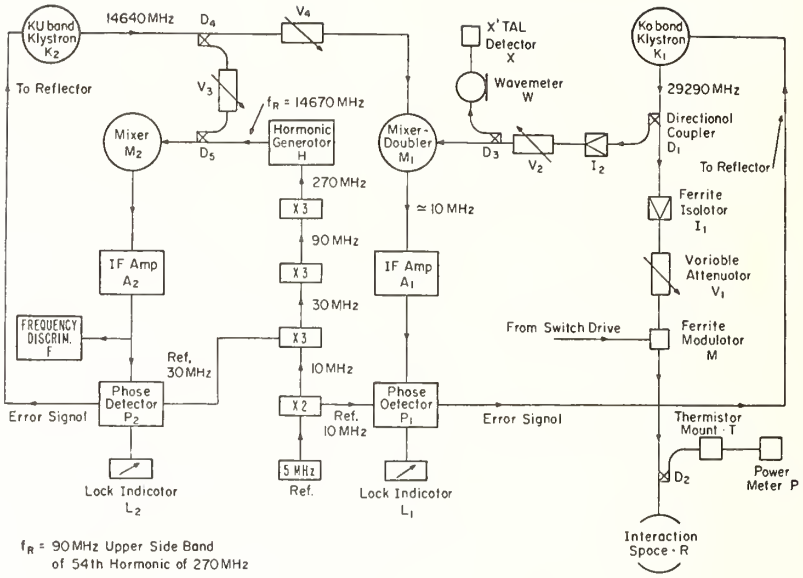


FIGURE 8. *Block diagram of the radiofrequency apparatus showing the double phase lock loop process for frequency stabilization.*

sensitive detector-amplifier. The resonance signal strength has been measured as a function of magnetic field, from which the resonance line center corresponding to the applied rf frequency has been determined. Figure 4 shows the Zeeman transition frequency as a function of magnetic field, according to the theory of Zeeman effect [4], from which the measured $\alpha \rightarrow e$ resonance line center in magnetic field space has been extrapolated to yield the zero field Lamb shift value (8).

3. Experimental Apparatus

Figure 5 is a schematic of the main experimental setup. A broad division of the apparatus consists of six parts: (1) interaction space, (2) electrode structure, (3) rf apparatus, (4) magnetic field equipment, (5) detection system, and (6) helium supply.

The interaction of the α state ions with rf field (frequency 29290 MHz) takes place in the space contained between the plates of a spherical mirror Fabry-Perot resonator. Rf power is fed into the resonator through a single iris whose size has been chosen such that the product of the Q of the resonator ($\simeq$1400) and the percent power absorbed into the resonator (35 percent) is maximum for resonator separation to contain 8 half wavelengths. The rf quenching region is essentially defined by the ion beam thickness and the spot size of the rf power ($\simeq$1.24 cm).

The electrode structure (fig. 6) consists of a tungsten wire filament (diam 0.25 mm, length $\simeq$1.9 cm) supported under tension ($\simeq$4 lb) and maintained at 200 V negative with respect to the acceleration grid G.I. This value of 200 V for the electron acceleration has been used in all runs, and is chosen to optimize the rf-induced signal on the basis of excitation characteristic curves (fig. 7). The normal value for the electron bombarding beam current in all experimental runs is 4.4 mA. The total emission current from the filament is servocontrolled to a constant value.

The rf power required to induce the αe transition is provided by a K_a band klystron (K1) and is fed to the Fabry-Perot resonator after passage through a variable attenuator (V1) and a ferrite power modulator (M) (see fig. 8). Frequency stabilization of the rf source is achieved by a double phase-lock loop process with a stability better than 1 part in 10^7. Due attention was paid to power drift contributions from source, resonator, and power meter, so that the absolute power level in the resonator is stable to better than 0.1 percent during the period of a run.

The normally employed value for the power level is sufficient to give 63 percent quenching level, while 35 percent and 95 percent quenching levels are also used to evaluate the power-dependent systematics.

A magnet with pole caps tapered down to 28 cm in diameter and a gap length close to 65 mm provided the magnetic field. A proton resonance absorption apparatus with its rf oscillator covering the frequency range from 60 MHz to 75 MHz is used in field measurement. The width of the proton resonance signal is normally about 200 mG. The measured field inhomogeneity over the dimensions of the quenching region is less than 15 ppm. The field at the probe position is measured to be lower by 2500$\pm$100 Hz compared to the field at the center of the quenching region.

The detection system consists of three parts: (1) an aluminum filter (≃1700 Å thick) to reduce background due to visible and long wavelength uv while permitting a larger fraction of the 304 Å uv light to pass; (2) a glass disc coated with sodium salicylate scintillator used to convert the uv light to visible light (≃50 percent efficiency for 304 Å signal); and (3) a detector tube with appropriate spectral sensitivity (S-11), located in a soft iron housing, the output from which is fed to the phase sensitive detector-amplifier as well as a vacuum tube voltmeter.

Commercial helium purified after passage through cold charcoal traps is used in the experiment. The pressure at the input position to the system is measured by a capacitance manometer. The absolute pressure at the interaction space is obtained by interpolation. The normal value for the operating helium pressure is 4.6 mtorr, while results at about 2.3 mtorr and 8.5 mtorr are also obtained to yield the final zero pressure extrapolated result.

4. Line Shape

4.1. Panaromic Resonance Curves

The rf quenching level Q of the metastable α state ions at an incident power level to the resonator P is

$$Q=(I/I_0)\times 100=[1-\exp(-\Gamma_\alpha\tau_\alpha C_{\alpha e})]C_{eg} \quad (1)$$

where I is the signal level at the detector at the power level P, I_0 is the maximum available signal, τ_α is the α state ion exposure time to the rf field, $C_{\alpha e}$ and C_{eg} are correction factors for $\alpha\rightarrow e$ and $e\rightarrow\alpha'$, β' transition matrix element variation with magnetic field. Γ_α is the α state ion induced decay rate given by

$$\Gamma_\alpha=\frac{(S/h)\,(e^2/\hbar c)\,\Gamma_e\,|\,\langle\alpha\,|\,Z\,|\,e\rangle\,|^2}{[(\nu-\nu_0)^2+\mu^2]} \quad (2)$$

where S is the incident rf power in ergs/(cm^2-s) in the resonator (at the input power level P), Γ_e is the spontaneous decay rate of the e state ion,

$$|\,\langle\alpha\,|\,Z\,|\,e\rangle\,|^2=3a_0^2/4,$$

a_0 is the Bohr radius, ν is the frequency separation of the levels α and e, $\nu_0=\nu_{\rm rf}=29{,}290$ MHz and $\mu_0=\Gamma_e/4\pi=800$ MHz.

Without any loss in the shape of (2) we may replace ν in eq (2) by [4]

$$\nu(\nu_p)=(\nu_{\alpha e}-\mathfrak{G})=\tfrac{1}{2}(g_s-g_l)\,(\mu_0H/h)+\tfrac{1}{2}(\Delta E/h)\times[(1+\tfrac{2}{3}\chi+\chi^2)^{1/2}-1] \quad (3)$$

where

$$\chi=(g_s-g_l)\mu_0H/\Delta E; \quad (4)$$

μ_0 in eqs (3) and (4) represents the Bohr magneton

$$\mu_0H=2\pi\,(\mu_0/\mu_e)\,(\mu_eH/\gamma_pH)\nu_p \quad (5)$$

where μ_e/μ_0 is the electron magnetic moment in Bohr magnetons [5] and $\mu_eH/\gamma_pHh=\omega_e/\omega_p$, the ratio of cyclotron frequency of a free electron to that of a proton in the same magnetic field H [6]; g_s and g_l are the spin and orbital g values for the electron corrected for anomalous part and nuclear motion respectively and ΔE is the $2^2P_{\frac{3}{2}}$–$2^2P_{\frac{1}{2}}$ fine structure energy level separation. Thus it should be noted that $\nu(\nu_p)$ is the frequency separation of levels α and e at the magnetic field H for which the proton resonance frequency is ν_p, over and above the zero field $\alpha\rightarrow e$ level separation.

Once the resonance line center magnetic field value is determined, the appropriate value for $\nu(\nu_p)$ is calculated through eq (3) and the zero field value for the Lamb shift is obtained from

$$\nu_0(\nu_p)=\nu_0-S \quad (6)$$

where $\nu_0=\nu_{\rm rf}$.

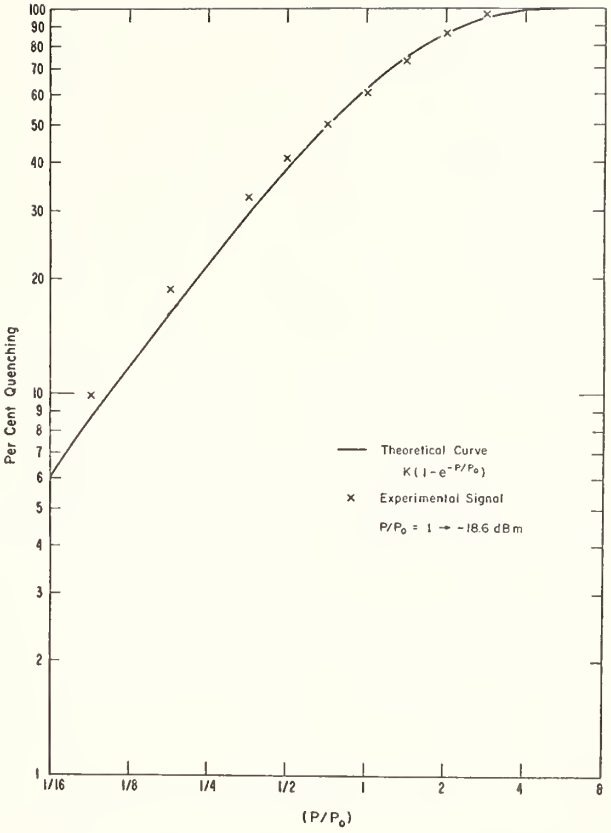


FIGURE 9. RF *quenching curve for the metastable ions at 4.6m mtorr for* He *pressure.*
Experimental quenching values are adjusted to fit the theoretical curve. P_0 is that power level for which $Q=63$ percent.

Figure 9 shows the percent rf quenching signal level as a function of rf power level at normal values for electron beam current and helium pressure at the approximate resonance center field value. Such curves are also used at other operating pressure values to obtain the power value required for 63 percent quenching. Other power levels used at the normal operating helium pressure (4.6 mtorr) are 35 percent and 95 percent quenching levels.

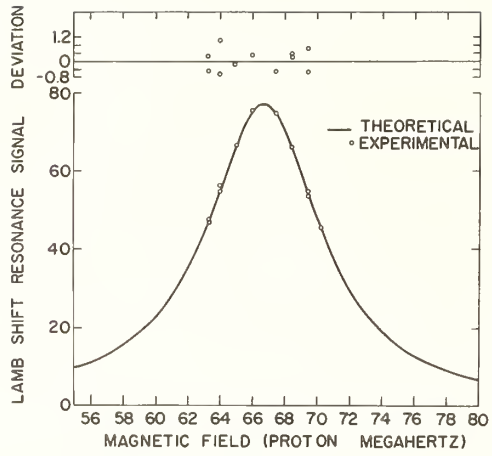


FIGURE 10. *Panoramic resonance curve for the αe transition resonance signal as a function of magnetic field.*
Electron beam current = 4.4 mA, helium pressure = 4.6 mtorr, rf quenching 63 percent. At the top deviation of experimental results from calculated values are shown on an expanded scale.

Under low power approximation, eq (1) reduces to

$$I = KC_{\alpha e}C_{e0}/[(\nu(\nu_p) - \nu_0')^2 + \mu^2] \qquad (7)$$

where

$$\nu_0' = \nu_{\mathrm{rf}} - \mathcal{S}$$

and

$$K = (I_0 S/h)\,(e^2/\hbar c)\,\Gamma_e\,|\,\langle \alpha \,|\, Z \,|\, e\rangle\,|^2. \qquad (8)$$

We used the line-shape expression [7] in analyzing our signal measurements at all operating conditions. The necessary corrections for deviations from the low power approximation are made later. The procedure used to obtain data in a single run consists of signal measurements at eight selected magnetic field values, in four pairs, each pair approximately symmetric on either side of the resonance. Their values lie in the range from 63.30 to 70.15 proton megahertz, and the field inhomogeneity in this range of fields is the same as described in section 3. The scheme of travel among these fields is zig-zag, and is also such that the initial pair of fields is repeated at the end of the run to check for the stability of the system during the period of the run.

Figure 10 shows a typical Lamb shift resonance line at the normal operating conditions of the apparatus. The width of the line is in close agreement with our calculation, and is 1000 MHz.

4.2. Ion Removal Processes

The time for which an α state ion is exposed to the rf field is different for different ions; it depends on the position of the ion formation, the velocity at this position, and the potential distribution along the path length of the beam. Since the ion dynamics in this problem are complicated, we make the simple constant velocity approximation under which $\tau_\alpha = d_{\mathrm{eff}}/\mathrm{v}$ where v is an average ion velocity. The d_{eff} is obtained by approximating the gaussian spatial variation of Γ_α along the magnetic field direction by a rectangular function having the same peak value and area as the true function; actually, d_{eff} is the width of this rectangular function ($\simeq$1.59 cm).

The various processes that govern the motion of ions and, hence, influence τ_α are (1) natural decay ($\tau_\alpha(N) = 1.9$ ms); (2) ion thermal and recoil motions ($\tau_\alpha(TR) = 16.5\ \mu$s); (3) space charge effects, where space charge electric fields give rise to a potential variation along the path of the beam ($\tau_\alpha(SP) = 0.276\ \mu$s); (4) Stark quenching, where motional and space charge electric fields cause mixing of α state ions with adjacent short-lived P states ($\tau_\alpha(ST) = 1.64\ \mu$s); (5) ion-atom collision quenching ($\tau_\alpha(IA) = 0.245\ \mu$s); and (6) other processes, such as nonradiative transfer of α state energy to an atomic electron ($\tau_\alpha(O)$). The values for τ_α in parenthesis are for 4.4 mA electron beam current and 4.6 mtorr He pressure. Treating the contributions from these sources as independent, the resultant value for τ_α is computed from

$$1/\tau_\alpha = 1/\tau_\alpha(N) + 1/\tau_\alpha(TR) + 1/\tau_\alpha(SP) + 1/\tau_\alpha(ST) + 1/\tau_\alpha(IA) + 1/\tau_\alpha(O) \qquad (9)$$

which gives for τ_α a value 0.126 μs.

The value computed for τ_α from $\Gamma_\alpha\tau_\alpha = 1$ for 63 percent quenching is $\simeq 0.11 \pm 0.02\ \mu$s. The agreement may be fortuitous, but excellent. Clearly, the longitudinal space charge electric fields play an important role in the ion removal. The magnetic field influence on the ion removal process has been considered in the estimates of corrections and uncertainties in the measured values for $\mathcal{S}$.

5. Corrections and Uncertainties

Effects influencing the observed line shape and the measured resonance line center are discussed in four categories. It should, however, be noted that the use of the zero pressure extrapolation method for the final result eliminates the need for part of the pressure-dependent corrections which are characteristic of an assemblage of ions or atoms.

5.1. Effects Intrinsic to Ion

(1) Shifts in α and e state energy levels: Such shifts caused by the presence of external magnetic field are properly considered in eq (7), requiring no separate correction. Shifts caused by the presence of electric fields are due to Stark effect. The transverse motional electric field due to thermal and recoil motions of the ion (21.6 V/cm) and transverse space charge electric field due to electron beam current (estimated $\simeq$39 V/cm) can couple α and f states and β and e states. The former component is a rotating vector in the xy plane whereas the latter is of variable magnitude at different positions in the x direction. Their sum is taken to be (8.8$\pm$45.6) V/cm leading to a correction of $+0.004 \pm 0.118$ MHz in $\mathcal{S}$.

(2) αe transition matrix element variation: The necessary correction factor $C_{\alpha e}$ for the magnetic field dependence of the αe transition matrix element

is included in eq (7), requiring no separate correction.

(3) Changes in final state density: This would lead to a field dependence of Γ_e, the $e \rightarrow \alpha', \beta'$ transition probability but since it is proportional to the third power of the $n=2 \rightarrow n=1$ energy separation, and since the fractional change in this energy separation over the working range of fields is negligible (a part in 10^8), no correction is required.

(4) Overlap correction: We infer from an observed αf transition a 6 percent σ component for the rf field in our apparatus, which excites βe and αf transitions. However, overlap from αf transition is weaker at higher quenching levels since a smaller number of α state ions are available to contribute, while the same is not true with β ions. The assigned corrections and uncertainties are, $\theta\delta\omega(\alpha f)+\delta\omega(\beta e)\pm(1/2)\delta\omega(\beta e)$ where $\theta=1$, 1/2, and 0 for the three quenching levels. For 35 percent, 63 percent, and 95 percent quenching power levels at 4.4 mA and 4.6 mtorr, these corrections are then -0.017 ± 0.058 MHz, $+0.096\pm0.113$ MHz, and $+1.231\pm0.616$ MHz, respectively, for S. Since S states in higher levels are not metastable (lifetime $\simeq 10^{-8}$ s), no overlap contributions are expected from these states.

(5) Signal intensity redistributions: The correction factor C_{eg} in eq (7) corrects for the magnetic field variation of the e state to ground state matrix element, requiring no separate correction.

(6) Rf Stark effect: From Lamb's analysis (HIII, p. 274) the shift in the angular frequency $\omega(\alpha e)$ is calculated according to

$$\Delta\nu = \Gamma_p \tau_\alpha^{-1} / [16\pi\omega(\alpha e)] \qquad (10)$$

where $\omega(\alpha e)=29290$ MHz, $\Gamma_p=10^{10}$/s, and for normal operational conditions, $1/\tau_\alpha=8.54\times10^6$/s. Assigning $1/(2\tau_\alpha)$ as the correction uncertainty, the necessary corrections are -0.060 ± 0.030 MHz, -0.117 ± 0.058 MHz and -0.345 ± 0.172 MHz to S at the three quenching levels, respectively.

5.2. Effects Characteristic of an Assemblage of Atoms or Ions

(1) Field dependence of available α state ions: Once equilibrium conditions are established in the system, the production rate of α state ions is essentially constant, since the magnetic field does not influence either the electron energy or the electron path lengths.

(2) Stark effect due to ionic space charge fields: At the normal operating conditions of our apparatus, the estimated values for the transverse and longitudinal electric fields (8.8 ± 45.6 V/cm and 12.5 ± 12.5 V/cm, respectively) would have required for the Stark effect, a correction of $+0.090\pm0.252$ MHz to S but for the zero pressure extrapolation procedure.

(3) Field dependence of space charge motion and collision quenching: We have used the Helm, Spangenberg, Field [7] formalism and charge conservation principle to estimate $\tau_\alpha(SP)$ in section 4.2. According to this formalism, although the ion beam is compressed in the transverse direction at a higher magnetic field, the ion removal time is not affected, since the ion charge density and ion production rate have similar dependence on the ion beam thickness. The longitudinal space charge electric field is independent of the beam compression effect and this results in a magnetic field independent collision quenching cross-section. We assign an uncertainty of ±1.0 MHz for these effects while we continue to look for better analysis of this problem to reduce such a large uncertainty.

(4) Field dependence of resonance radiation absorption: At the existing ion densities in our apparatus, we estimate the resonance radiation reabsorption to be an insignificant process [3].

5.3. Deviations From Ideal Apparatus

(1) Magnetic field inhomogeneity: The correction required for the field difference between the position of probe and interaction space is -0.570 ± 0.025 MHz.

(2) Rf power variation: The magnetic field dependence of the various rf equipment components due to residual field variation requires a correction of -0.050 ± 0.050 MHz.

(3) Grid transmission variation: The magnetic field dependence of the electron bombarding current due to GI transmission variation requires a correction of -0.250 ± 0.050 MHz.

(4) Photomultiplier gain variation: This gain varies by 0.7 percent over the working range of fields, but the data obtained are corrected for this prior to processing. The associated uncertainty is ±0.160 MHz.

(5) Field dependence of quantum conversion: The magnetic field dependence of fluorescent radiation flux from sodium salicylate converter [9] requires a correction of -0.125 ± 0.125 MHz.

(6) Filament bending effect: A group of runs have been made (No. 3, table I) with a filament deliberately displaced by 1 mm at center in order to enhance such an effect. In view of the good internal agreement of the results of this run with other precision groups, no separate correction has been made.

(7) Field effects on associated electronic apparatus: The residual field variation at the position of electronic equipment has been found to require a correction of -0.125 ± 0.125 MHz.

5.4. Corrections for Method of Analysis

(1) Least-squares calculation: The signal amplitude measurements at the eight selected field values have been least-squares fitted to the line shape expression [7], using a three parameter fit, with the signal resonance center value, resonance center amplitude, and resonance line width as the parameters, from which the parameter values, as well as their standard deviations, are obtained. The results at various helium pressure values are then weighted-least-squares fitted to a straight line, where weighting

TABLE 1. *Precision experimental results*

Group no.	Filament no.	No. of runs	Operating conditions			
			He pressure (mtorr)	Beam current (mA)	R.F. level % Q	$\nu_0 \pm \Delta\nu_0$ (MHz)
I	1	13	2.4	4.4	63	15 244.657±0.395
II	2	7	2.5	4.4	63	15 245.059±0.700
III	1	9	4.7	4.4	63	15 244.243±0.547
IV	1	9	8.3	4.4	63	15 244.125±0.629
V	3	8	4.8	4.4	55 & 82	15 243.423±1.203
VI	4	6	4.6	4.4	35	15 243.230±0.975
VII	4	6	4.6	4.4	95	15 244.652±2.002
VIII	4	6	4.6	4.4	35/95	15 243.821±1.560

Weighted mean values for resonance centers of each group are listed under $\nu_0 \pm \Delta\nu_0$. The mean result of group V with displaced filament agrees well with other groups. Groups V, VI, VII used a lower transmission filter resulting in lower signal to noise. In group VII power broadening contributes to further reduction of signal to noise. Precision results are based on groups I, II, III and IV which are used to obtain zero pressure extrapolated result. In groups V, VI and VII quenching power level is the variable. Group VIII shows results of normalization procedure where the ratio signals (signal with 35% quenching/signal with 95% quenching) are least square fitted to ratio signal expression obtained from (7), where observed signals were corrected for known rf power effects prior to calculating ratios.

coefficients are obtained from the inverse squares of the standard deviations of individual resonance center values. The statistical error in the zero pressure interpolated value for S thus obtained is ±0.411 MHz whereas the pressure gradient and its standard deviation are (0.108±0.081 MHz)/mtorr.

(2) Postulated, modified Lorentzian line shape: Data processing has been done by the use of low-power approximated line shape (eq (7)), whereas at 63 percent quenching level the actual line shape is given by eq (1). In order that the αe matrix element variation, the Zeeman curvature and the ion removal time variation are (as indicated in eqs 9 and 1) properly corrected for, a data processing method correction +2.328±0.267 MHz to S is required.

6. Results and Discussion

The zero pressure extrapolated result for the resonance center obtained from groups I to IV of table 1 is:

$$\nu_0(\nu_p) = \nu_{rf} - S = 15244.947 \text{ MHz}. \qquad (11)$$

The slope of the pressure extrapolation procedure indicates a correction of −0.497±0.372 MHz for S if all data were taken at the single pressure of 4.6 mtorr. However, this slope might not include the pressure dependent effects to the full extent due to the fact that our lowest operating pressure is not sufficiently low and our consistent choice of 63 percent quenching level complicates the analysis further.

Table 2 shows necessary corrections made to the zero pressure extrapolated result for S obtained with 63 percent quenching. The present result thus obtained for S with three standard deviations for statistical error plus systematics is

$$S(He^+)_{n=2} = 14046.2 \pm 2.0 \text{ MHz}. \qquad (12)$$

We are presently examining better ways of analyzing the systematics due to τ_α variation. This result for S (eq 12) agrees with all previous experimental results, although the agreement with that of Lipworth and Novick [3] is only within three standard deviations. This result agrees with the most recent theoretical value [10]

$$S(He^+)_{n=2} = 14044.5 \pm 5.2 \text{ MHz}. \qquad (13)$$

TABLE 2. *Corrections and uncertainties to the Lamb shift with 63 percent quenching power level (in MHz)*

1. Effects intrinsic to ion:	
(1) Stark effect	+0.004±0.118
(2) Overlap effect	+0.096±0.113
(3) Rf Stark effect	−0.117±0.058
2. Effects characteristic of ion assemblage:	
Field dependence of space charge motion[a]	±1.000
3. Deviations from ideal apparatus:	
(1) Magnetic field inhomogeneity	−0.570±0.025
(2) Rf power variation	−0.050±0.050
(3) Grid transmission variation	−0.250±0.050
(4) Detector gain variation	±0.160
(5) Field dependence of quantum convertor	−0.125±0.125
(6) Field effects on electronic equipment	−0.125±0.125
4. Correction from method of analysis	
Data processing method[a]	+2.328±0.267
Total	+1.191±0.803
5. Statistical error	±0.411

[a] These two effects have correlated uncertainties.

7. References

[1] The series of papers on experiments performed by Lamb and his collaborators have been referred to in the literature as H I, H II, H III, H IV, H V and H VI. These are
H I: Lamb, W. E., Jr., and Retherford, R. C., Phys. Rev. **79**, 549 (1950).
H II: Lamb, W. E., Jr., and Retherford, R. C., ibid. **81**, 222 (1951).
H III: W. E. Lamb, Jr., ibid. **85**, 259 (1952).
H IV: Lamb, W. E., Jr., and Retherford, R. C., ibid. **86**, 1014 (1953).
H V: Triebwasser, S., Dayhoff, E. S., and Lamb, W. E., Jr., ibid. **89**, 98 (1953).
H VI: Dayhoff, E. S., Triebwasser, S., and Lamb, W. E., ibid. **89**, 106 (1953).

[2] Taylor, B. N., Parker, W. H., and Langenberg, D. N., The Fundamental Constants and Quantum Electrodynamics, Academic Press, New York (1969). See also Rev. Mod. Phys. **41**, 375 (1969).

[3] The series of papers on He^+ are:
He^+ I: Lamb, W. E., Jr., and Skinner, Miriam, Phys. Rev. **78**, 539 (1950).
He^+ II: Yergin, P. F., Ph.D. thesis, Columbia University (unpublished).
He^+ III: Novick, R., Lipworth, E., and Yergin, P. F., Phys. Rev. **100**, 1153 (1955).
He^+ IV: Lipworth, E., and Novick, R., Phys. Rev. **108**, 1434 (1957).

[4] Bethe, H., and Salpeter, E. E., Quantum Mechanics of One and Two Electron Systems, Academic Press, New York (1957); See also II III.

[5] Wilkinson, D. T., and Crane, H. R., Phys. Rev. **130**, 852 (1963).

[6] Lambe, E. B. D., thesis, Princeton University (1955).

[7] Helm, R., Spangenberg, K., and Field, L. M., Electrical Communication **24**, No. 1 (Mar. 1947).

[8] Lipworth, E., Ph.D. thesis, Columbia University (1955).

[9] Narasimham, Musti A., J. Opt. Soc. Am. **59**, 781 (1969).

[10] Appelquist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970).

DISCUSSION

R. A. Brown: Could you say a few more words about the rf Stark effect?

M. A. Narasimham: Yes. We have used Lamb's analysis in Hydrogen III from which the rf Stark effect is proportional to 1 over tau, where tau is the lifetime of the ion. Now, we have both estimated tau from the various ion removal processes that govern the lifetime of the ions in the interaction space, as well as computed it from the amount of power that it takes in order to induce 63 percent quenching level, because at that power level you have gamma alpha tau alpha equal to 1. So these two are in very good agreement. We use the tau alpha we got from there, and with one over tau alpha, plus or minus one half as much has been put in as an uncertainty in the correction. This comes straight from the analysis of Lamb. I can give you the correct equation from my notes if you would be interested in it.

R. A. Brown: This is the Bloch-Siegert effect?

M. A. Narasimham: The Bloch-Siegert effect is due to the oscillating magnetic field component but not the electric field component.

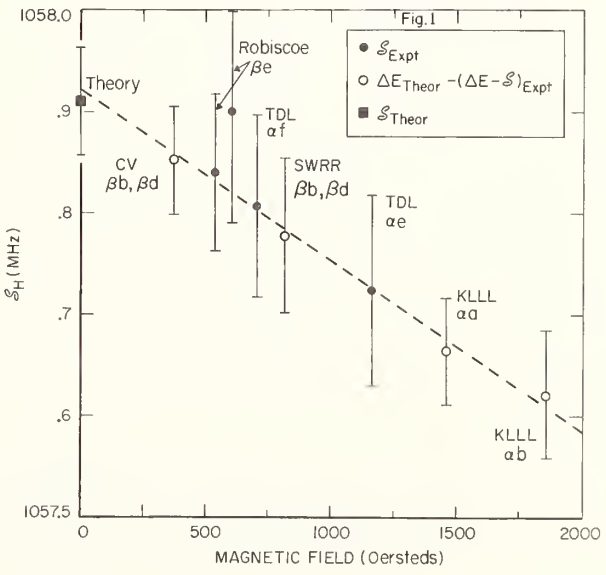
Editor's Note: At this point, H. Metcalf showed without comment a slide due to S. L. Kaufman. It seemed to us to be of sufficient interest to be included in the Proceedings. S. L. Kaufman has agreed and encloses the following comment to accompany the slide.

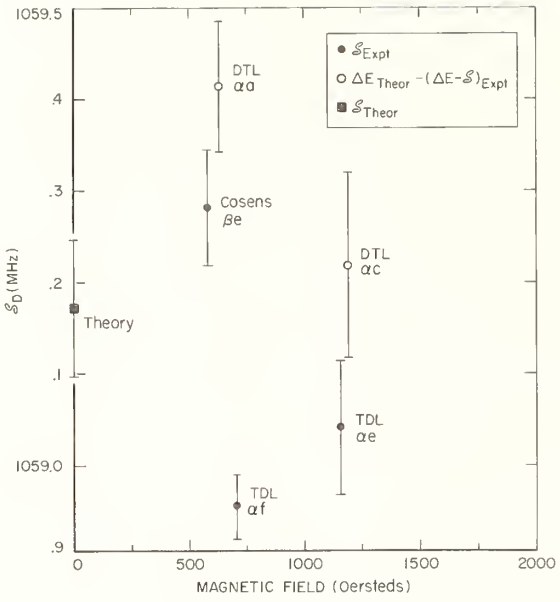
Comment on Hydrogen Fine Structure Results

S. L. Kaufman

Several measurements have now been made of the $2^2P_{3/2}$–$2^2S_{1/2}$ interval ($\Delta E - \mathcal{S}$) in hydrogen. Discrepancies outside the statistical errors exist between some of the measured values. Our purpose is to show that these discrepancies seem to be systematic in a way not inconsistent with measured values of the Lamb shift $\mathcal{S} = 2^2S_{1/2} - 2^2P_{1/2}$. We present the effect without interpretation, assuming that its source is experimental, and hoping to encourage the search for its cause.

We note simply that the experiments have all been performed in different magnetic fields, and that the results seem to depend weakly on the field. In order to put $\mathcal{S}$ on the same graph with ($\Delta E - \mathcal{S}$), we take the $2^2P_{3/2}$–$2^2P_{1/2}$ interval ΔE to have the value 10969.026(42) MHz as given by Taylor et al. This yields a value of the Lamb shift for each measured value of ($\Delta E - \mathcal{S}$). These values of the Lamb shift, together with those obtained from the more direct measurements, are now plotted against the magnetic field values at which the transitions were observed (fig. 1). The theoretical value of Applequist and Brodsky is shown at zero field. The linear relationship is striking, as shown by the (not fitted) straight line.





The first warning about this graph is that the error bars represent only one standard deviation. Also, for the theoretical value, one-third the "limit of error" is shown. Thus the entire effect could be a statistical accident.

Note that the two highest-field values were measured using the "bottle" method, while all of the remaining values come from atomic-beam measurements. Our effect could thus be a systematic error influencing, say, only the "bottle" measurements.

A further warning against taking the effect too seriously appears when the same graph is plotted for deuterium (fig. 2). There is clearly no sign of a systematic effect in this case, regardless of the value of ΔE assumed.

EDITORS' ADDITIONAL NOTE: The Robiscoe values of $\mathcal{S}_H$ do not include his recent +0.040 MHz kinematic correction.

S. J. BRODSKY: I'll take comments if anyone would like to ask a question.

D. KLEPPNER: Can you explain it? (*Laughter*)

Precision Measurement of the Fine Structure and Lifetimes of the $(1s2s2p)^4P_J$ States of He^- and Li* [1]

Robert Novick and David Weinflash

Department of Physics, Columbia University, New York, N. Y. 10027

A series of measurements has been completed and others are in progress for the precise determination of the energies and lifetimes of the metastable autoionizing $(1s2s2p)^4P_J$ states in He^- and Li*. Since these states have lifetimes in the range from 5 to 500 μs and since the fine and hyperfine structure intervals are of the order of a wave number, it will be possible to determine the splitting with a relative accuracy of one part in 10^6 or better. Present measurements in ^{6}Li and ^{7}Li are at the level of one part in 10^4 and a preliminary crude estimate has been made of the $^4P_{5/2}$–$^4P_{3/2}$ interval in He^-. Further work is in progress with radio-frequency techniques to obtain very much more precise results. Considerable theoretical effort must be made before these results can be interpreted in terms of fundamental constants. However, it is clear that if the necessary numerical wave functions can be obtained with sufficient accuracy, and if the various radiative and relativistic corrections can be evaluated, then these results will provide a precise new independent value for the Sommerfeld fine structure constant.

Key words: Autoionization; metastable atoms; Zeeman quenching.

1. Introduction

The metastable autoionizing $(1s2s2p)^4P_J$ states of He^- and Li* provide an unusual opportunity to check the accuracy of quantum calculations for three-electron systems. The fine-structure intervals and lifetimes of these states are a sensitive test of our understanding of electron correlation effects and the competition between various spin-dependent interactions. In addition, the Fermi contact hyperfine interaction of the unpaired s-electrons provides a critical test of our ability to calculate the electron density at the nucleus. If numerical wave functions are obtained with sufficient accuracy and the various radiative and relativistic corrections are evaluated, very precise measurements of the level splittings will provide an independent value for the Sommerfeld fine-structure constant, α, accurate to six places or better. Alternatively, if we assume that α is known from other work, then such precise experimental values could be used either as a new test of quantum electrodynamics or to place limits on new interaction terms that may be postulated for the Hamiltonian of simple atomic systems [1].

The feature of these states which lends them to high precision measurement is their metastability: they are stable against electric-dipole emission and decay only by spin-dependent autoionization to a ground-state configuration plus a free electron. The lifetimes are in the range from 5 to 500 μs, implying natural widths as low as 2 kHz. The resolution of an rf resonance will be limited only by field homogeneity and the time spent in the rf region, which may be as long as 100 μs. Since the levels are separated by $\sim$1 cm^{-1}, it will be possible to determine the levels with a relative accuracy of one part in 10^6 or better.

A further important property of these states is that each level of the ^4P_J multiplet has a different lifetime. Hence, differential populations are spontaneously created in flight, both between J-levels, and in an axial static magnetic field, between M_J-levels (Zeeman quenching). This effect can be exploited for state selection and the production of polarized particles.

2. Metastable Autoionizing States

The earliest interest in atomic states metastable against both autoionization and radiative decay concerned attempts to understand the existence of the helium negative ion, first observed by Hilby [2]. The possibility of this ion existing in the lithium-like $(1s^22s)^2S_{1/2}$ ground-state configuration was rejected when variational calculations [3] failed to show a positive electron affinity.[2]

Ta-You Wu suggested [4] that the $(1s2s2p)^4P$ term of He^-, in which two of the electrons are excited, could be bound with respect to the $(1s2s)^3S$ term of neutral helium. Elaborate variational calculations have confirmed this conjecture, but at least 20 terms are required in the wave function to produce a

[1] Work supported by National Aeronautics and Space Administration under Grant NSG-360 and by National Science Foundation, Grant NSFGP-13749.

[2] G. J. Schulz has reported the observation of a resonance in the elastic scattering of electrons on helium at 0.45 eV which he attributes to the formation of a temporary negative ion with the configuration $(1s^22s)$. See *Proceedings of the Fourth International Conference on the Physics of Electronic and Atomic Collisions, Quebec, 1965*, edited by L. Kerwin and W. Fite (Science Bookcrafters, Inc., Hastings-on-Hudson, New York, 1965), p. 117.

positive electron affinity for helium in the $(1s2s)^3S$ state [5]. The selection rule $\Delta S=0$ forbids decay to the helium ground state plus a free electron by either electric-dipole emission or by autoionization via the Coulomb potential between the electrons. The state is therefore metastable and was postulated to account for the observations of He^-.

Analogous metastable autoionizing states were predicted for the alkalis. These quartet states are produced by the excitation of a single core electron and have energies much greater than the first ionization potential of the atom. Of particular interest in the present connection are the $(1s2s2p)^4P_J$ states of lithium which have exactly the same configuration as predicted for He^-. The $\Delta S=0$ selection rule makes Li*, like He^-, metastable with a lifetime of a few μs.

It was realized [6] that the breakdown of LS coupling in both He^- and Li* would mix all but the $^4P_{5/2}$ state with ^2P states which autoionize very rapidly (lifetime on the order of one Bohr period: $\sim 10^{-14}$ s). Thus even a very small doublet admixture would strongly quench the $^4P_{3/2}$ and $^4P_{1/2}$ states. Finally, Kroll pointed out that even $^4P_{5/2}$ could decay by the tensor part of the spin-spin interaction. The $^4P_{3/2}$ and $^4P_{1/2}$ states can also decay directly by spin-dependent interactions as well as indirectly via the doublets, leading to generally shorter lifetimes. In the case of Li*, the metastable state has an energy of 57 eV, or more than ten times the first ionization energy of Li, yet its lifetime is about 6 μs.

The term "differential metastability" has been coined to describe the existence of distinct lifetimes for each J-state of the multiplet. A further proliferation of lifetimes occurs in an external magnetic field and/or in the presence of hyperfine structure.

In the exploratory work carried out so far, the splittings of the $^4P_{J,F}$ states are obtained by diagonalizing a semiphenomenological Hamiltonian which includes fine and hyperfine structure terms as well as the Zeeman interaction:

$$\begin{aligned} H = {} & c_{SO}\mathbf{L}\cdot\mathbf{S} \\ & + c_{SS}[3(\mathbf{L}\cdot\mathbf{S})^2 + 3/2(\mathbf{L}\cdot\mathbf{S}) - L(L+1)S(S+1)] \\ & + a_c\mathbf{I}\cdot\mathbf{S} + \mu_0 H(g_L L_z + g_S L_z + g_I I_z). \end{aligned} \quad (1)$$

Magnetic dipole and electric quadrupole hyperfine terms were also included in the analysis, but were not varied in fitting data [7]. In the absence of hyperfine structure, let the eigenstates of given M_J in an external magnetic field be indexed by the superscript i:

$$|\Psi^i{}_M(H)\rangle = \sum_J a^i{}_{JM}(H)\,|J, M\rangle. \quad (2)$$

If the separation between the levels is large compared to their width, the decay rate of $|\Psi^i{}_M(H)\rangle$ is well-defined and given by a weighted sum of the zero-field decay rates γ_J of the multiplet:

$$\gamma^i{}_M(H) = \sum_J |a^i{}_{JM}(H)|^2 \gamma_J. \quad (3)$$

Expressions of the same form as eqs (2) and (3), with M replaced by F, give the eigenstates and decay rates of the hyperfine levels with no external field. The state of highest F ($F_{max}=5/2+I$) is a pure $J=5/2$ state and thus, in general, is longest-lived. With both hyperfine and external fields, the summation in eq (3) extends over J and F to obtain the decay rates for states with a given M_F. The most important consequence of the mixing by an external field is to increase the decay rate of the $^4P_{5/2,F}$ states, thereby "Zeeman-quenching" the metastable signal reaching the detector in an atomic beam experiment. However, the state of highest F and $|M_F|$ is not mixed and retains the pure $^4P_{5/2}$ deeay rate, $\gamma_{5/2}$.

3. Lithium

The early work on lithium and the other alkalis established the existence of autoionizing states with lifetimes in the μs range [8]. A schematic diagram of

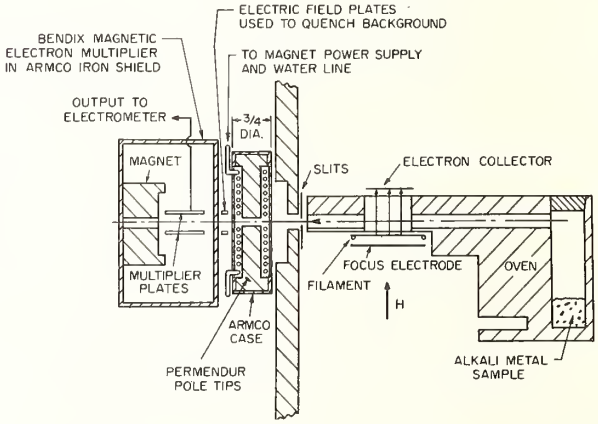


FIGURE 1. *Schematic diagram of the first Zeeman quenching atomic beam apparatus (ref. [8]).*

the apparatus used in these investigations is shown in figure 1. A beam of metastable atoms was produced by electron bombardment of lithium atoms emanating from an oven. The metastables were detected by collecting the ion-electron pairs that result from autoionization. The original detector was simply a pair of parallel metal plates with a small potential across them to enhance collection efficiency. Greater sensitivity was achieved by replacing the metal plates with a magnetic strip electron multiplier. To prevent charged particles from entering the detector, a honeycomb grid was placed across the entrance. The threshold in lithium for the exeitation of the metastable signal was 57.3 ± 0.3 V. Since the creation of the metastable involves a spin flip, the excitation function is characteristic of exchange collisions: the cross section attains its maximum within a few eV of the threshold energy and then decreases very rapidly. An auxiliary detector downstream from the main detector was sensitive to photons and

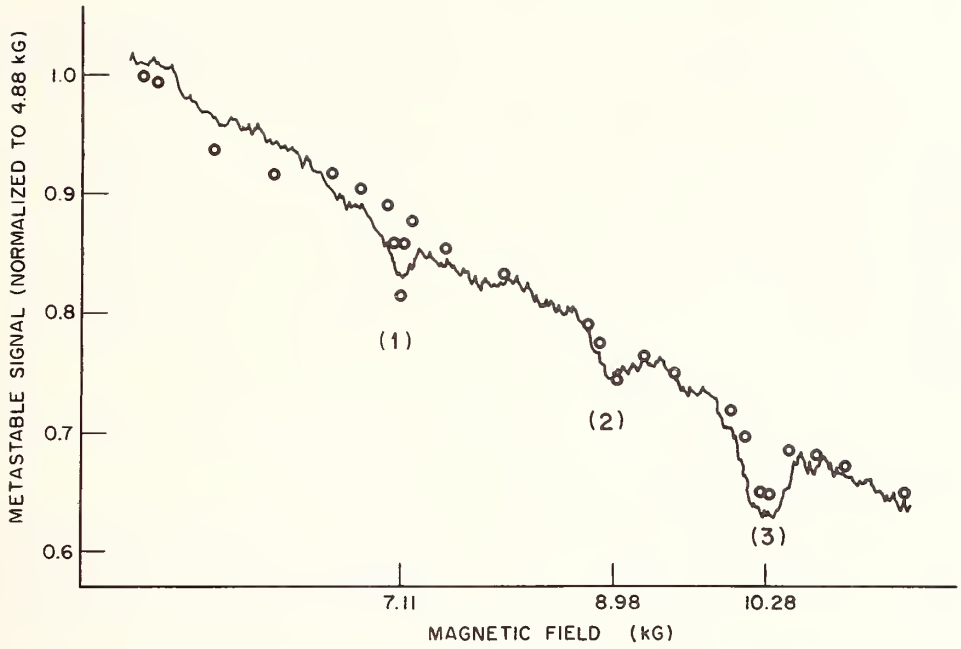


FIGURE 2. *Typical data for* ^{7}Li *showing dips in the Zeeman quenching curve.*
These and similar features in ^{6}Li were studied in greater detail by signal-averaging techniques (ref. [7]).

metastables but not the ion-electron pairs, and showed no marked increase in signal strength at the metastable threshold.

The metastable detector was movable over a distance of 3 cm and the lifetime of the lithium metastable, which was assumed to be pure $^4P_{5/2}$ by the time it reached the detector, was determined to be 5.1 ± 1.0 μs.

Quenching of the metastable signal by electric and magnetic fields was also observed. The magnetic quenching curves of ^{7}Li and ^{6}Li showed several resonance-like dips which were attributed to strong mixing of short-lived and long-lived states at level anticrossings. The quenching curves for both isotopes were studied in greater detail with the entire apparatus (source and detector) enclosed in a uniform magnetic field to prevent nonadiabatic transitions in the fringing region [7]. Typical data for ^{7}Li are reproduced in figure 2. Three anticrossing dips in ^{7}Li were investigated and ascribed to long-lived $J\approx5/2$ levels anticrossing short-lived $J\approx1/2$ levels. The anticrossing levels are illustrated in figure 3. In ^{6}Li, three anticrossings of $J\approx5/2$ states were found: two involving $J\approx3/2$ states and one involving a $J\approx1/2$ state. The centers of the dips were accurately located, and the widths measured, by signal-averaging techniques.

A theoretical analysis showed that in the vicinity of the anticrossings, a two-state treatment of the levels is adequate. The fine-structure intervals $^4P_{5/2}$–$^4P_{3/2}$ ($\equiv\Delta_{53}$) and $^4P_{5/2}$–$^4P_{1/2}$ ($\equiv\Delta_{51}$) were determined to a few parts in 10^4 by finding, and demon-

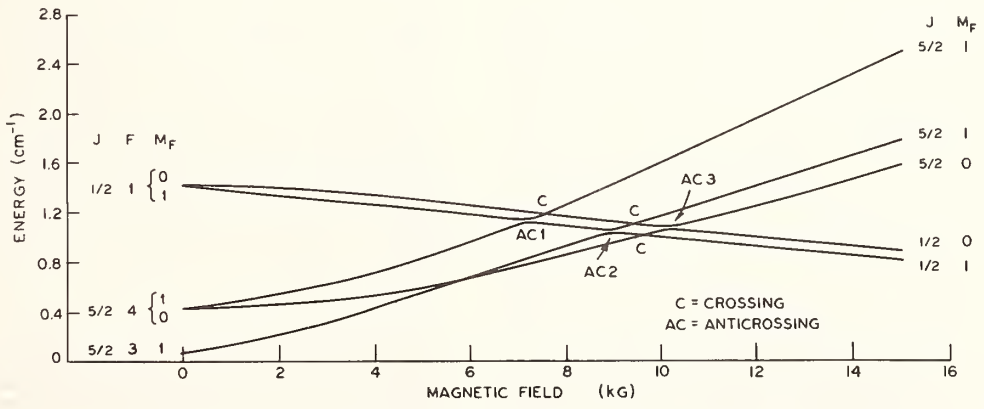


FIGURE 3. *Anticrossings in* ^{7}Li *energy diagram responsible for dips in figure 2 (ref. [9]).*

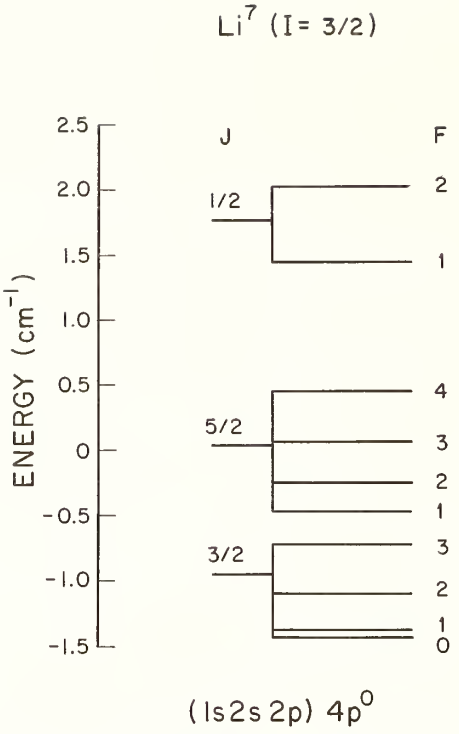


FIGURE 4. *Experimental level diagram for* ^{7}Li (*ref.* [7]).

strating the uniqueness of, a set of values for c_{SO}, c_{SS}, and a_c in eq (1) that gave anticrossings in ^{6}Li and ^{7}Li at the experimentally observed fields. The $^4P_{5/2}$ and $^4P_{1/2}$ lifetimes were determined by the widths of the corresponding anticrossings. Finally, the $^4P_{5/2}$ lifetime was reevaluated, taking into account the considerable hyperfine mixing, which raised the value about 16 percent [9]. The results for the energies of the $(1s2s2p)^4P_{J,F}$ states of ^{7}Li appear in figure 4. The lifetimes and structure constants are given in table 1.

This work provided the first definitive assignments for lines observed in the course of a study of the Li II spectrum [10] which had no place in the normal spectra of the atom or ion. It had been suggested that these lines arose from transitions to the metastable $(1s2s2p)^4P_{J,F}$ ground quartet-state from higher quartets. Assuming the upper state to be $(1s2s3s)^4S$, the new experimental results for the level structure of $(1s2s2p)^4P_{J,F}$ in ^{7}Li were used to calculate a theoretical line shape for the unassigned 2934 Å line. The calculated line profile is in excellent agreement with the spectroscopic data [7]. This agreement strongly supports the term assignments and the structure of the ^4P_J states.

The level scheme described above predicts the location of *crossings* between states of different M_F to within 1 G. These states can be coupled in the neighborhood of the crossings by an appropriate rf magnetic field oscillating perpendicular to the static field H_0, and transitions detected by a drop in metastable signal. Under favorable experimental

TABLE 1. *Fine-structure and lifetime of* $(1s2s2p)^4P_J$ *in* Li*
(Intervals in mK = 10^{-3} cm^{-1}; lifetimes in μs)

Quantity	Experiment[a]	Theory[b]
$\Delta_{53}(^4P_{5/2}-^4P_{3/2})$	+997.34±0.66	+575
$\Delta_{51}(^4P_{5/2}-^4P_{1/2})$	−1724.70±0.54	−2610
c_{SO}	−154.47±0.30	−325
c_{SS}	+184.47±0.14	+185
$a_c(^7Li)$	+172.09±1.12	+172
$a_c(^6Li)$	+65.16±0.42	+65
$\tau_{5/2}$	5.8±1.2	5.88
$\tau_{3/2}$	0.46±0.10	0.30
$\tau_{1/2}$	0.14±0.07	>10.0*

* Involved large cancellation due to interference

[a] References [7, 9].
[b] Reference [14].

circumstances, the line-width is determined only by the natural width of the short-lived state, which is ~0.25 MHz for $J \approx 3/2$ and ~0.75 MHz for $J \approx 1/2$. Determination of the fine and hyperfine structure could be about 200 times more accurate than in the static quenching experiment, and precision of a few parts in 10^6 or better could be achieved. Tentative identification of a 30 MHz resonance in ^{7}Li at 5374 G has been made [11]. This is roughly 3 G below the crossing of a $J \approx 3/2$, F = 1, $M_F = 0$ level with a $J \approx 5/2$, F = 1, $M_F = -1$ level, as calculated from the level scheme of figure 4. The resonance has proved to be elusive and work is continuing on this and other predicted hyperfine resonances.

4. The Helium Negative Ion

Results on the helium negative ion, which prompted the whole line of research, are as yet primitive, at least for the fine structure. The negative ion of the ^{4}He isotope (I = 0) has been studied by time-of-flight techniques in an axial magnetic field and Zeeman quenching observed at fields as low as 400 G. Lifetimes of roughly 500, 10, and 16 μs have been obtained for the $J = 5/2$, 3/2, and 1/2 states, respectively.[3] In addition, the $^4P_{5/2}$–$^4P_{3/2}$ fine-structure interval has been estimated to be 0.036±0.009 cm^{-1} from the Zeeman quenching data between 350 and 1400 G [12].

Figure 5 is a schematic diagram of the apparatus used in the study of He^-. A beam of the negative ion is prepared by double charge exchange, a technique proposed by Donnally [13]. A 3000-eV He^+ beam is extracted from an 80 MHz rf discharge and passes through a potassium vapor where it suffers two collisions with neutral potassium atoms. The first collision produces the metastable $(1s2s)^3S$ state of

[3] These lifetimes are in apparent disagreement with an 18.2 μs lifetime for He^- reported by Nicholas, Trowbridge, and Allen, Phys. Rev. **167**, 38 (1968). However, these investigators did not consider the differential metastability of He^- and based their conclusions on a maximum flight time of 1 μs during which they observed a ~5 percent decay of their beam. This agrees with the decay observed during the first microsecond of decay in our work when both the long-lived and short-lived components are present in the beam.

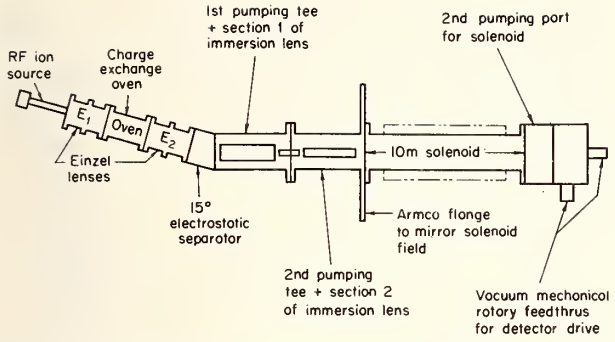


FIGURE 5. *Schematic diagram of the* He^- *source and detection regions (ref. [12]).*

He^0. The second collision results in another charge exchange and produces the desired He^- state. The potassium vapor density can be varied to maximize the production of He^-, and a negative ion current of 100 nA is easily obtained. The beam emerging from the potassium vapor is electrostatically separated into its charge components and the total positive beam is collected as a monitor on beam stability. The He^- beam is deflected into a series of electrostatic lenses and decelerated to any desired energy above ~50 eV. The beam then passes into a 10-m drift tube wound to provide axial magnetic fields up to 1500 G. The ion source and detection regions are separated by a differential-pumping aperture. The pressure on the source side is 3×10^{-6} torr when helium gas is admitted to the rf discharge, while the pressure in the detection region is at least an order of magnitude lower.

The He^- beam is detected by a Faraday cup which is movable in vacuum nearly the full length of the drift tube. A high transparency grid, electrically insulated from the Faraday cup, is placed across the entrance of the detector. When the negative retarding voltage on the grid is greater than 20 V, the 19.7-eV electrons produced by autoionization in flight of He^- cannot enter the detector and the current recorded is due to He^- alone.

The dependence of the He^- metastable beam intensity on the experimental variables (distance from source, magnetic field, beam velocity, and residual gas pressure) was analyzed to extract the lifetimes and fine structure of the $(1s2s2p)^4P_J$ states of He^-. The bulk of the data was taken at a beam energy of 100 eV and a field of 400 G. The plot of metastable current versus distance is fit in all cases to 1 percent by a weighted sum of two exponential components. The closest position of the detector to the source corresponds to a flight time of 7 μs for a 100-eV beam. If the observed weights are extrapolated back to the source, the intensities at production are roughly equal. The decay rate γ_S of the short-lived component was essentially independent of field, velocity, and pressure. The decay rate γ_L of the long-lived component increased with all these variables, and was extrapolated to zero residual pressure using data at various pressures and velocities. The long-lived component was identified with a weighted average of the $^4P_{5/2}$ substates, while the short-lived component was assumed to be an average of the $^4P_{3/2}$ and $^4P_{1/2}$ states. This assignment is supported by the agreement between the relative strengths of the two components with their statistical weights and the increase of the decay rate of the long-lived component with magnetic field.

More detailed information on the lifetimes and fine-structure intervals was obtained directly from the Zeeman quenching curves. The field dependence of the metastable current (shown in fig. 6) yielded values for the $^4P_{5/2}$–$^4P_{3/2}$ fine-structure interval Δ_{53} and the difference between the $^4P_{3/2}$ and $^4P_{5/2}$ decay rates, $\gamma_{3/2}-\gamma_{5/2}$. With these two atomic parameters fixed, all three zero-field decay rates were determined from the decay-length data. In figure 7, the data are fit with three exponential components. The weights extrapolated back to the source at S are proportional to the statistical weights of the assigned J-values. Only the $^4P_{5/2}$–$^4P_{1/2}$ fine-structure interval, Δ_{51}, remains inaccessible to the present experimental approach. Table 2 gives the preliminary lifetimes and fine-structure obtained in this work.

There are probably no narrow anticrossings in the negative ion of the isotope ^3He (I=1/2) such as those observed in ^6Li and ^7Li. Manson's [14] theoretical level diagram for the ^3He ion is shown in figure 8. The hyperfine splitting is calculated to be considerably larger than the Δ_{53} fine-structure interval and comparable to the Δ_{51} interval. Narrow anticrossings occur between states which would in

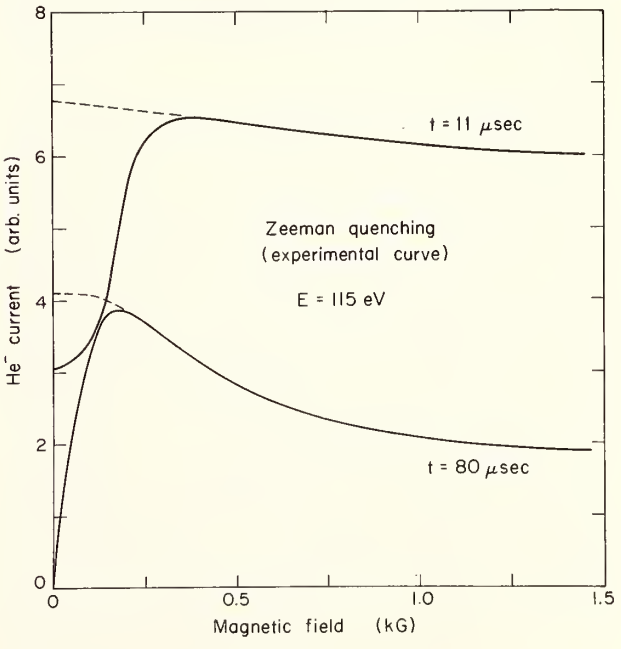


FIGURE 6. *Zeeman quenching of* He^- *at two flight times.* The metastable current drops at low fields because of solid angle losses incurred when the beam is not contained by a magnetic field. The dashed lines represent extrapolation assuming quadratic dependence on the field.

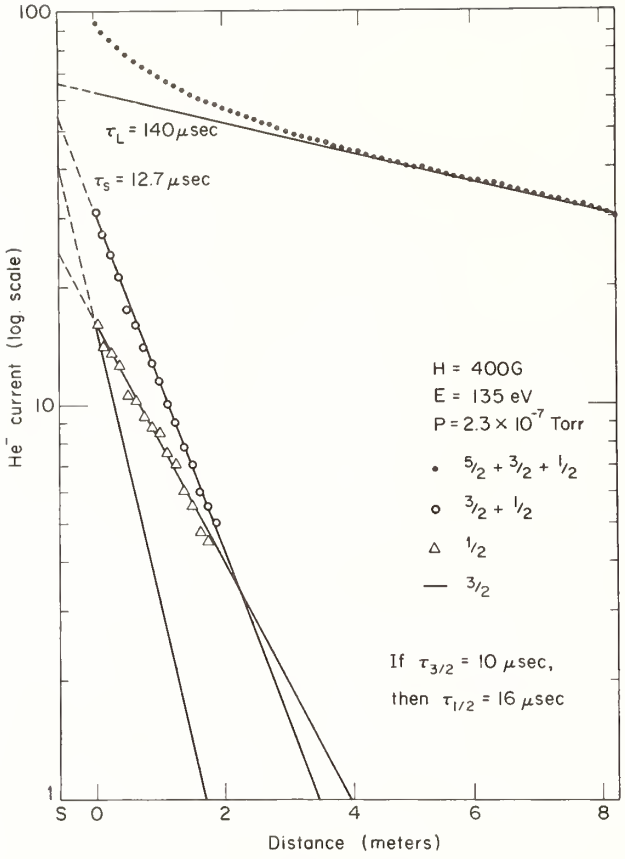


FIGURE 7. He^- *time-of-flight data fitted with two and then three exponential components.*
The two-component fit (circles and dots) is based on these data alone. The three-component fit uses $\tau_{3/2}$ from Zeeman quenching (fig. 6) to determine $\tau_{1/2}$ (triangles). The relative strengths of the three components extrapolated back to the source at S agree with the statistical weights of the assigned J-states.

fact cross if they were not weakly coupled by a perturbation small compared to the Hamiltonian which defines them. There are *broad* Zeeman anticrossings in the energy levels for ions of both He isotopes, but these are of little experimental interest and despite the smallness of Δ_{53}, the lowest of these occurs around 3700 G, still too high for the present apparatus. It would be prohibitive to produce a magnetic field more than a few kilogauss over a distance sufficient to appreciably quench the long-lived He^- states. These states mix with "short-lived" states that have lifetimes ~ 10 μs, or decay lengths of ~ 70 cm for a 100-eV beam.

Although not expected to exhibit anticrossings, the Zeeman quenching curve of the ^3He ion could give a value for Δ_{51}, the atomic parameter unmeasured by the work to date. Denote as $F=1'$ and $F=2'$ the hyperfine states with $J \approx 3/2$ to distinguish them from the states $F=1$, $J \approx 1/2$ and $F=2$, $J \approx 5/2$. Because $\Delta_{53}<0$, we would expect $\Delta_{32'}<0$ and $\Delta_{21'}<0$ in figure 8. However, Manson's calculation gives $\Delta_{21'}>0$; i.e., $F=2$ is higher then $F=1'$ because the latter is depressed by repulsion from $F=1$. This depression of the $F=1'$ level is sensitive to the

TABLE 2. *Fine-structure and lifetimes of* $(1s2s2p)^4P_J$ *in* $^4He^-$
(Intervals in $mK = 10^{-3}$ cm^{-1}; lifetimes in μs)

Quantity	Experimental[a]	Theory[b]
$\Delta_{53}(^4P_{5/2}-{}^4P_{3/2})$	±36±9	−68
$\Delta_{51}(^4P_{5/2}-{}^4P_{3/2})$	[c] −263	−314
c_{SO}	−49	−61.3
c_{SS}	11.4	11.4
a_c	—	−101
$\tau_{5/2}$	500±[d]200	1000
$\tau_{3/2}$	10±2	33
$\tau_{1/2}$	16±4	[e] 3500

[a] Reference [12].
[b] Reference [14].
[c] There is only experimental data for $|\Delta_{53}|$; the rest of experimental column uses $\Delta_{53}=-36$ and theoretical c_{SS} to calculate Δ_{51} and c_{SO}.
[d] See footnote 3.
[e] Involved large cancellation due to interference.

value of Δ_{51}. It is possible that Δ_{51} is actually too large to cause inversion of $F=1'$ and $F=2$, as Manson claims. If $\Delta_{21'}$ is significantly different from $\Delta_{32'}$, very precise Zeeman quenching data for the ^3He ion could determine the absolute value of $\Delta_{21'}$, although unfortunately not its sign. If the actual splittings are favorable, a value for Δ_{51} could be obtained by this method.

Very sharp rf resonances could be obtained at low frequency in the neighborhood of level *crossings* in the ion of either isotope. There are four $\Delta M_J=-1$ crossings in the ^3He ion between 400 and 700 G suitable for low-frequency π-transitions. They occur where $|J=5/2, M+1\rangle$ intersects $|J=3/2, M\rangle$, for

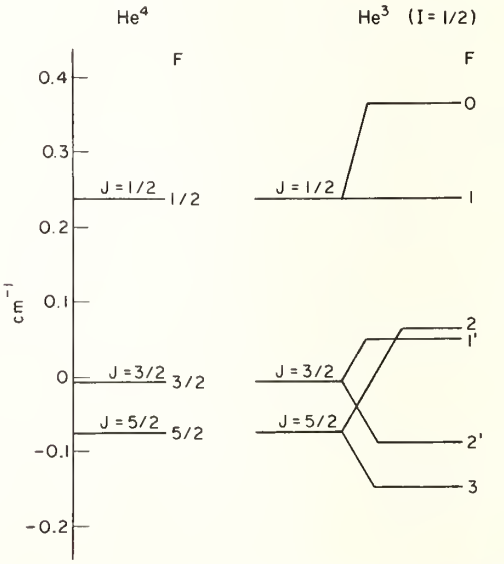


FIGURE 8. *Calculated energy level diagrams for the helium negative ion of the isotopes* ^4He *and* ^3He (*ref.* [14]).
A preliminary value (ref. [12]) for the $^4P_{5/2}$–$^4P_{3/2}$ splitting is approximately half the calculated value (see table 2). The ^3He states marked 1′ and 2′ have $J=3/2$ weight of 90 percent and hence are longer-lived than the states marked 2 and 3.

$-3/2 \leq M \leq +3/2$. The location of the crossing with $M_{3/2} = +3/2$ depends only on Δ_{53}; the location of the crossing with $M_{3/2} = -3/2$ is influenced strongly by Δ_{51}. Similar crossings occur in the ^{3}He ion between $F=3$ and $F=2'$ and between $F=2$ and $F=1'$. The contributions to the line-width are: (1) rf power level, which can be made low once the resonance is located; (2) time spent by the ion in the rf field, which can be as long as 100 μs; (3) inhomogeneities in the static magnetic field; (4) natural decay width of the short-lived state, which is $\sim$100 kHz.

At such time as resonance techniques are successfully applied to this problem there will be a number of radio-frequency and microwave transitions that can be readily studied. If we conservatively estimate that the line center can be determined to 1/10 of its width, we can measure the intervals in He^- and Li* to one part in 10^6 or better. This is an order of magnitude more precise than results for systems that are not metastable. For example, the 2^3P_1–2^3P_2 fine-structure interval in the $(1s2p)^3P_J$ state of neutral helium was measured by optical level crossing to 40 ppm [15]. The line-width of 2.7 G, equivalent to several MHz, arose from the radiative lifetime ($\sim 10^{-7}$ s) of the allowed-electric-dipole transition to $(1s2s)^3S$. A more recent determination of this interval by an atomic-beam optical-microwave resonance method is accurate to 3 ppm [16]. However, this precision is based on fitting the data to a theoretical line shape and locating the center of the line to 1.2×10^{-3} of its width. In hydrogen, the line-width of the 2^2P state is about 100 MHz, or roughly 1 percent of the $^2P_{3/2}$–$^2P_{1/2}$ fine-structure separation.

5. Comparison to Theoretical Work

As might be expected, the agreement between theory and experiment is better for Li* than for the weakly-bound He^-. Several calculations of the $(1s2s2p)^4P$ term energy in lithium are within the experimental limits of 57.3 ± 0.3 eV above the $(1s^22s)$ ground state (see Table II of ref. 8). Manson's values in lithium for the $^4P_{5/2}$ lifetime [17], the spin-spin constant c_{SS}, and the hyperfine constant a_c [14] are all in good agreement with experiment. The only serious discrepancy is in the value of the spin-orbit constant c_{SO}; the theoretical value is roughly twice the experimental value. In the $(1s2s2p)^4P_J$ states, c_{SO} includes a negative contribution from the magnetic field produced by the p-electron on the unpaired spins of the s-electrons, called the spin-other-orbit interaction. This overpowers the usual spin-orbit energy due to the p-electron orbit producing a magnetic field on its own spin; the net orbital interaction makes c_{SO} negative. Because c_{SO} is the difference of two comparable quantities, it is very sensitive to electron correlation effects, the exact form of the wave function, etc.

The situation is more difficult for He^-. The largest electron affinity obtained for He by variational calculations is Weiss' value of 0.069 eV [4], which is about 14 percent smaller than the experimental value of 0.080 ± 0.002 eV measured via a laser photodetachment technique by Brehm, Gusinow, and Hall [18]. Since the energy is the quantity most accurately calculated by variational techniques, the other quantities evaluated using these wave functions are in serious doubt.

The preliminary experimental value for the Δ_{53} fine-structure splitting is about half the theoretical splitting. If the theoretical value for c_{SS} is assumed to be accurate, as is suggested by the agreement in Li*, the experimental value for Δ_{53} can be used to estimate values for c_{SO} and Δ_{51} (see table 2).

Theoretical values for the lifetime of $^4P_{5/2}$ in He^- ranging from 266 to 1700 μs can be found in the literature [19, 20]. Most recent calculations include electron correlation in the initial state through a multiterm variational wave function. The most sophisticated calculation considers the effects of polarization of the $(1s^2)$ final state by the free electron and the consequent distortion of the continuum wave function [21]. Because the $^4P_{5/2}$ state can decay only via the tensor part of the spin-spin interaction, there is no ambiguity concerning the operator in the matrix element of the transition. A reliable and precise experimental value for the lifetime is therefore a clearcut test of the initial and final state wave functions and should inspire further theoretical efforts.

Future experimental efforts will aim at finding and precisely measuring the rf and microwave resonances described above. If the anticipated precision is obtained, the results will provide rigorous constraints on theoretical discussions of the three-electron system. Theoretical groups are considering the possibility of calculating very precise numerical nonrelativistic wave functions for three-electron states [22]. Clearly such results will be essential to the interpretation of the high precision experimental studies.

6. References

[1] Sakitt, B., and Feinberg, G., Phys. Rev. **151**, 1341 (1966).
[2] Hilby, J. W., Ann. Physik **34**, 473 (1939).
[3] Wu, Ta-You, Phil. Mag. **22**, 837 (1936).
[4] Wu, Ta-You, Phys. Rev. **58**, 1114 (1940).
[5] Holøien, E., and Geltman, S., Phys. Rev. **153**, 81 (1967). This reference contains previously unpublished work of A. W. Weiss.
[6] Holøien, E., and Midtdal, J., Proc. Phys. Soc. (London) **A68**, 815 (1955); **90**, 883 (1967).
[7] Feldman, P., Levitt, M., and Novick, R., Phys. Rev. Letters **21**, 331 (1968).
[8] Feldman, P., and Novick, R., Phys. Rev. Letters **11**, 278 (1963); "Atomic Collision Processes," edited by M. R. C. McDowell (North-Holland Publishing Company, Amsterdam, 1964), pp. 201–210; Phys. Rev. **160**, 143 (1967).
[9] Feldman, P., Levitt, M., and Novick, R., Phys. Rev. A **3**, 130 (1971).
[10] Herzberg, G., and Moore, H. R., Can. J. Phys. **37**, 1293 (1959).
[11] Levitt, M., Novick, R., and Skwire, S., "Columbia Radiation Laboratory Progress Report No. 19" (1969), page 6.

[12] Blau, L. M., Novick, R., and Weinflash, D., Phys. Rev. Letters **24**, 1268 (1970). New data have been obtained on He^- and the data in this paper reanalyzed.
[13] Donnally, B. L., and Thoeming, G., Phys. Rev. **159**, 87 (1967).
[14] Manson, S. T., Phys. Rev. (to be published).
[15] Colegrove, F. D., Franken, P. A., Lewis, R. R., and Sands, R. H., Phys. Rev. Letters **3**, 420 (1959).
[16] Pichanick, F. M. J., Swift, R. D., Johnson, C. E., and Hughes, V. W., Phys. Rev. **169**, 55 (1968).
[17] Manson, S. T., Phys. Rev. **145**, 35 (1966); Phys. Rev. Letters **23**, 315 (1966).
[18] Brehm, B., Gusinow, M. A., and Hall, J. L., Phys. Rev. Letters **19**, 737 (1967).
[19] Pietenpol, J. L., Phys. Rev. Letters **7**, 64 (1961).
[20] Laughlin, C., and Stewart, A. L., J. Phys. B: Proc. Phys. Soc. (London) **1**, 151 (1968).
[21] Estberg, G. N., and LaBahn, R. W., Phys. Letters **28A**, 420 (1968); Phys. Rev. Letters **24**, 1265 (1970).
[22] Jacobs, Verne (private communication).

The Proton Moment in Bohr Magnetons—and All That

Daniel Kleppner

Department of Physics and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass. 02139

The gamesmanship of fundamental constants is explained. The tactics are elucidated by the presentation of a new value for the ratio of the proton moment to the electron moment in the ground state of hydrogen. Measurements were made with an atomic hydrogen maser. The result is $-\mu_e(\mathrm{H})/\mu_p(\mathrm{H}) = 658.210\ 706\ 1(65)$.

Key words: Atomic hydrogen maser; proton magnetic moment.

1. Foreword

The fact that there is enough interest in fundamental constants to muster an International Conference suggests that the topic has come of age as a mature field of physics. Like all mature fields it has its own rules and conventions, and I think that it might be useful to review some of these before getting lost in the nitty-gritty of the work at hand.

To make a contribution to the field of fundamental constants one merely has to measure something simple as accurately as possible. In spite of the mystique that surrounds precision measurement the rules of the game are straightforward: the goal is to acquire credible digits (or *significant figures*, as they are known), the more the better. Whoever finds the most significant figures for a particular constant wins the local championship. Sportsmanship requires that the championship be relinquished gracefully if the competition should top your number of significant figures, and that it be forfeited instantly should you have the misfortune to lose any of your figures, the most grievous loss being to have the competition find that some of your figures are wrong.[1]

Unfortunately, preoccupation with significant figures frequently blinds participants to other aspects of the field, in particular to the distinction between *significant* figures, *interesting* figures and *useful* figures.

By *interesting* figures, we mean those digits which evoke more than the usual ho-hum awarded to a new measurement. For example, the magnetic moment of the proton is 2.792782(17) nuclear magnetons. Here we have six significant figures[2] (the seventh digit, in spite of its location, is not particularly significant) and one interesting figure, namely the first. As with all interesting digits, this one has an interesting history: When Otto Stern measured the proton moment in the early 1930's he was advised not to bother—elementary theory proved that the result would be one nuclear magneton. Fortunately Stern had a healthy disregard for elementary theory. His measurement, which involved the deflection of molecular hydrogen, was crude—he obtained only a single figure. However, this figure turned out to be *three* nuclear magnetons rather than one, and, as everybody agreed, that was interesting. It still is. Incidentally, though the interesting digits are found generally towards the beginning of a number, while the significant digits are found towards the end, this is not always the case; when they coincide great legends are born. Nevertheless, as a rule of thumb the more interesting digits are likely to be the least significant.

By *useful* figures we mean those digits which are neither significant nor interesting but are handy to have around. For instance, the tenth digit of $\Delta\nu(\mathrm{H})$, the hyperfine frequency of hydrogen, happens to be "1." This is not particularly interesting since theory runs into trouble back near the sixth digit, nor is it particularly significant since $\Delta\nu(\mathrm{H})$ can be measured to twelve digits (or at least eleven and one-half—a slight tendency to exaggerate is endemic to the field). However, it is a useful digit since anyone who needs to know $\Delta\nu(\mathrm{H})$ to ten figures is mighty glad that it has been measured to twelve. He has no need to worry, in contrast to the unfortunate soul who really

[1] This is only a bare outline of the rules of the game. We forgo describing such ploys as stealing the championship from yourself, thereby simultaneously forfeiting and winning it, or referring to one of your experiments which you would rather forget since it turned out all wrong not by your own name but by the name of a reviewer who had the misfortune to include your results in his compilation of the best values. These are not for the novice.

[2] Logically the first of a string of digits is the most significant. However, since it is always the last digit which causes all the work, cognoscenti regard the last figure as the most significant. This causes no confusion as long as you bear in mind that the last significant digit is the most important (and vice versa).

needs all twelve digits and who had better worry plenty. From the point of view of actually using a fundamental constant, the most significant figure is the least desirable.

Custom requires that every new significant figure be accompanied by a written estimate or guarantee of its reliability. (This is responsible for most of the personal injuries in the game, but apparently the custom is too well established to be changed.) It has been suggested that a notation be adopted which similarly indicates the interesting and useful figures. Thus interesting figures might be followed by exclamation points and useful figures could be heavily underlined to emphasize their workhorse properties. Appealing as this seems it is not really practical since it invites further value judgments, for instance an x to indicate that a scandal was attached to the digit (lower case for experiment, upper case for theory), a question mark to indicate that one doesn't believe the digit, or an arrow to indicate that as far as that digit is concerned you would rather fight than switch. The system would undoubtedly collapse from over-elaboration. Personally, I opt for the tradition of simply listing the digits as they come, leaving it to the reader's skill, ingenuity and experience to read between the digits.

2. The Role of the Proton Moment in Bohr Magnetons

To turn to the work at hand, we offer some new results for one of the less glamorous fundamental constants, the value of the magnetic moment of the proton in Bohr magnetons, μ_p/μ_B. There was a flurry of excitement over this quantity a few years ago when it was suspected that an error in its value might account for the hydrogen hyperfine structure discrepancy. (This, in faet, was the chief rationale for the experiment we describe below.) The hyperfine discrepancy vanished with the revision of the fine structure constant, and so, alas, though we report here some new significant figures, we cannot claim any new interesting figures. Nevertheless, there is hope; accurate knowledge of μ_p/μ_B can help provide a highly sensitive check on the theory of diamagnetic shielding. It may lead the way to investigation of high order effects of binding on nuclear moments. Furthermore, since nuclear magnetic resonance of the proton provides one of the simplest calibration methods for magnetic fields, the ratio μ_p/μ_B is frequently needed as a transfer standard to link experimental observations with better known fundamental constants. A case in point is its role in obtaining the fine structure constant from a Josephson measurement of $2e/h$ [1]:

$$\alpha^{-1}=\left(\frac{c}{4R_\infty\gamma_p}\frac{\mu_p}{\mu_B}\frac{2e}{h}\right)^{1/2}.$$

What is really needed in this expression is the gyromagnetic ratio of the electron, $\gamma_e=e/mc$. This has not been measured accurately, so one introduces the gyromagnetic ratio of the proton and then converts using $\gamma_e=(\mu_B/\mu_p)\gamma_p$. Finally, μ_p/μ_B is useful as one of the pieces of stochastic input data in the compilation of the table of fundamental constants.

Ideally, we should compare the magnetic moments of a free electron and free proton. Since nobody has figured out how to do this we must settle for various combinations of second best. Briefly, these consist of comparison of a magnetic resonance transition of electrons, either free or bound in the hydrogen atom, with the magnetic resonance of protons, either in a molecule or in the hydrogen atom. Klein [2] has compared the free electron with the proton in water, and Lambe [3] has compared the electron in hydrogen with the proton in water. These experiments, and in fact the entire status of the subject up to mid 1969, have been reviewed by Taylor, Parker, and Langenberg in their magnum opus [4]. Hence, we devote the rest of this paper to some results we have obtained since then.

3. The Experiment

We have determined the ratio of the electron moment to the proton moment, $\mu_j(\mathrm{H})/\mu_p(\mathrm{H})$, (or, alternatively, the ratio of their g factors, $g_j(\mathrm{H})/g_p(\mathrm{H})$) as observed in atomic hydrogen. To obtain the ratio of intrinsic moments for the free particles it is necessary to correct for binding and other effects. Fortunately, hydrogen is so simple that these can be calculated to high accuracy, as reported by Grotch and Hegstrom elsewhere in these Proceedings.

The Hamiltonian for ground state hydrogen in an applied magnetic field H_0 is

$$\mathcal{H}=ha\mathbf{I}\cdot\mathbf{J}-g_j\mu_B\mathbf{J}\cdot\mathbf{H}_0-g_p\mu_B\mathbf{I}\cdot\mathbf{H}_0.$$

Three unknowns are involved: the hyperfine constant a, and g_j and g_p. To determine these completely, three frequencies must be measured. Fortunately, the hyperfine frequency is well known, so in principle only two frequencies remain to be measured. In practice, we measured all three frequencies indicated in figure 1. ν_e, 9192 MHz, corresponds to an "electron" transition: it depends dominantly on the interaction of the electron with the magnetic field. ν_{p1}, 644 MHz, and ν_{p2}, 776 MHz, are "proton" transitions: they depend dominantly on the hyperfine interaction, but also slightly on the interaction of the proton with the applied field.

The method involves using an atomic beam maser similar to a conventional hydrogen maser except that it operates in an applied field of 3500 G. The radiating atoms are confined by a 2 cm diameter PTFE lined bulb in a microwave cavity. The electron transition is detected by observing the free precession of the electron magnetization following a 90° pulse. Simultaneously, a proton transition is observed by double resonance. A continuous signal is applied to the proton; when resonance is achieved the electron radiation is quenched. Further experimental description is given in an account of an earlier version of this experiment which was reported

several years ago [5]. Details of the present work will be published elsewhere.

The strong coupling of the proton to the electron results in a serious loss of sensitivity in the experiment. As a result, the proton transition frequencies have to be determined fractionally 40 times more precisely than the resulting magnetic moment ratio. Thus, to determine μ_p/μ_B to 1 part in 10^8, ν_p has to be determined to 2.5 parts in 10^{10}. The basic problem is the familiar problem of resonance physics: to find the center of a resonance curve of unknown shape by analysis with an inexact theory which is too complicated to apply rigorously.

The line width of the electron transition was typically 100 Hz. Since ν_e had to be measured only to 100 Hz, there was no difficulty in accomplishing this. Furthermore, the narrow line width meant that

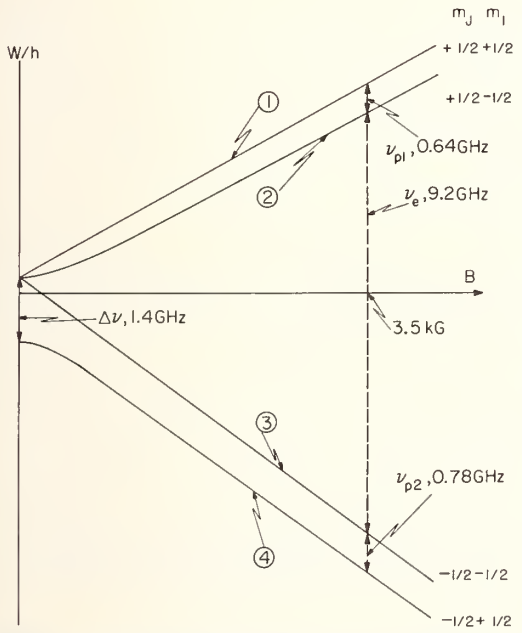


FIGURE 1. *Energy levels of ground state hydrogen in an applied magnetic field.*
The observed transitions are indicated.

the direct effect of systematic errors such as cavity pulling could be neglected. However, indirect effects due to coupling of the electron and proton transitions were, as we shall describe, most troublesome.

The major source of the electron line width was due to inhomogeneity of the magnetic field. The inhomogeneities arose in approximately equal proportions from irregularities in the magnetic field and from the diamagnetism of the quartz storage bulb. Motional averaging of the field reduced the broadening by a factor of 25 from what would have been observed if the radiating atoms were stationary.

The burden of work lay in locating the center of the proton line. Figure 2 illustrates the method. The first frame shows the electron signal following a 90° pulse. (The oscillation within the deeay envelope is due to a frequency offset introduced for display.) Succeeding frames show the effect of a continuous signal near the proton transition. The electron signal is modulated, the maximum effect occurring when the proton signal is at resonance (last frame). The proton resonance curve is displayed by taking the difference of the radiated electron signal with and without the applied proton signal. Figure 3 compares an experimental resonance curve with a set of points calculated from the theory of the lineshape. A Lorentzian curve has been fitted to each. Aside from the slight peaking at the line center the Lorentzian

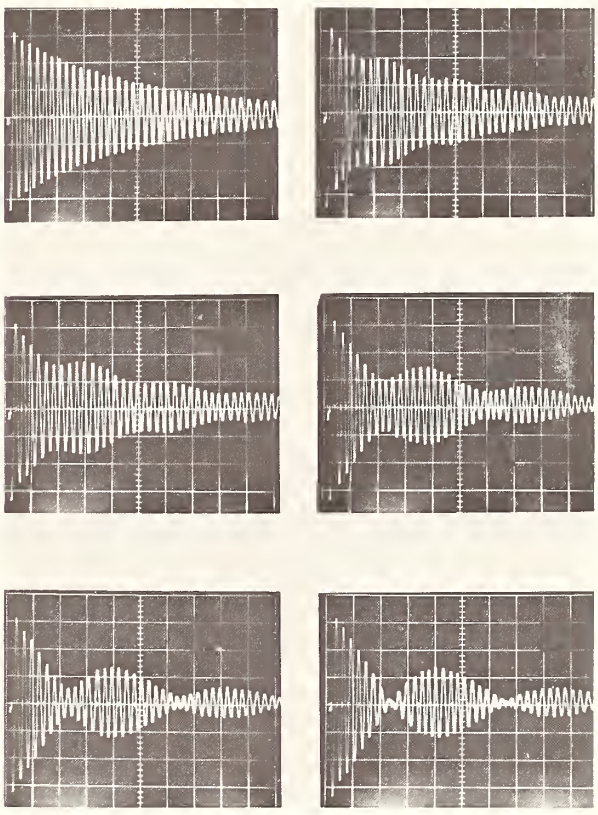


FIGURE 2. *Free precession signal from the electron in the presence of a signal near the proton resonance.*
The proton signal is swept progressively towards resonance from the upper left hand to lower right hand picture. Full scale is 0.01 s.

is a reasonably good fit, and since it is much easier to handle analytically than the exact shape, we used it in our analysis. (In defense it should be pointed out that the exact lineshape is distorted symmetrically and that this procedure does not introduce errors in finding the true line center.)

4. Systematic Errors

The major source of experimental error arose from frequency shifts associated with motional averaging. Briefly, if the resonance frequency of a radiating

atom is given by $\omega_0 = \gamma H_0$, then in the case of an ensemble of atoms which are confined to a small volume but which are free to move so that each atom makes many excursions across the volume during the radiation lifetime, one expects that the resonance is centered at the frequency $\omega_0 = \gamma \langle H_0 \rangle$, where $\langle H_0 \rangle$ is the space average magnetic field. As pointed out by Brenner [6], this is not necessarily true. There are a number of effects which shift ω_0, often by a large fraction of the line width. In particular, shifts can occur if the amplitude of the magnetic field varies asymmetrically or if the rf driving field is non-uniformly distributed. The magnitude of the shift increases with the gyromagnetic ratio, and though it was negligible for the proton transition it could easily be several Hz for the electron transition. Since the latter had to be determined only to 100 Hz this was of little direct consequence. However, the electron and proton resonances are not independent—they are coupled by a common level. As is well known for coupled oscillators, shifting the frequency of one mode pulls the frequency of the other mode. Fortunately, the degree of pulling depends on the line widths of the two modes, and this gave us a handle on the effect since the proton line width could be broadened at will by increasing the proton driving signal.

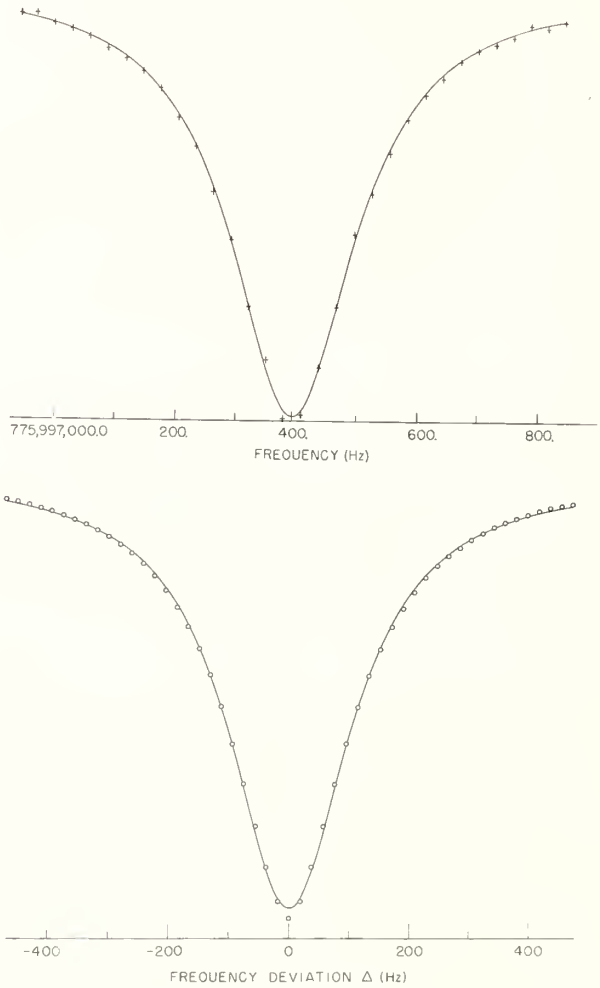


FIGURE 3. *Experimental data points (upper plot) and data points calculated by integrating the equations of motion (lower plot).*
The curves are Lorentzian.

Figure 4 shows measurements made under three different applied field gradients. The fitted curves are parabolas. Ideally, one could adjust the field gradients until the g-factor ratio was independent of the proton line width. This turned out to be impractical. Instead, our approach was to take data at various line widths and to extrapolate to zero line width. Such a procedure is almost guaranteed to cause grief unless one has a good understanding of the lineshape. In our case, this required knowing details of the magnetic field gradients, and this simply was not possible. Instead, our approach was to model the experiment on a computer with various field configurations, and to study the frequency shifts as a function of line widths. As one might expect, to first approximation the effect of motional averaging was to introduce small distortions in the resonance curve. For instance, in figure 5 we display a set of data points and a set of calculated points for the case of significant distortion due to motional

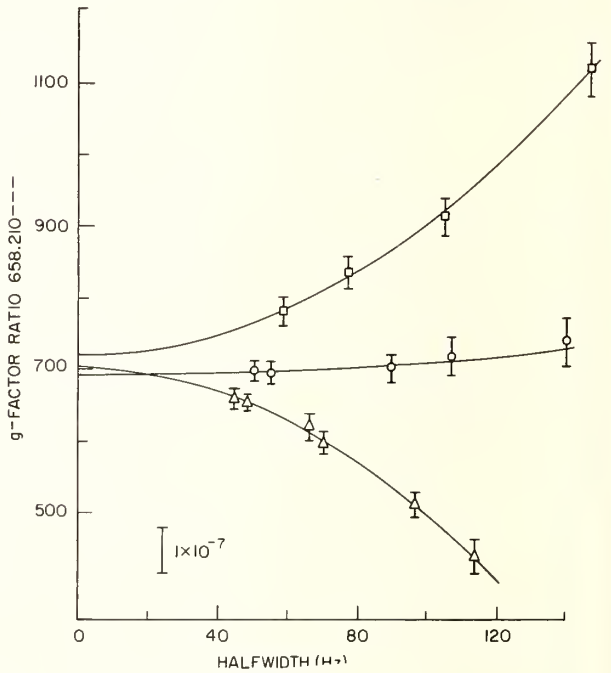


FIGURE 4. *Effects of field gradients on the g-factor ratio as determined for various proton line widths (the line width is varied by changing the driving power level).*
The curves are parabolas centered at zero linewidth. The three curves correspond to different applied magnetic field gradients.

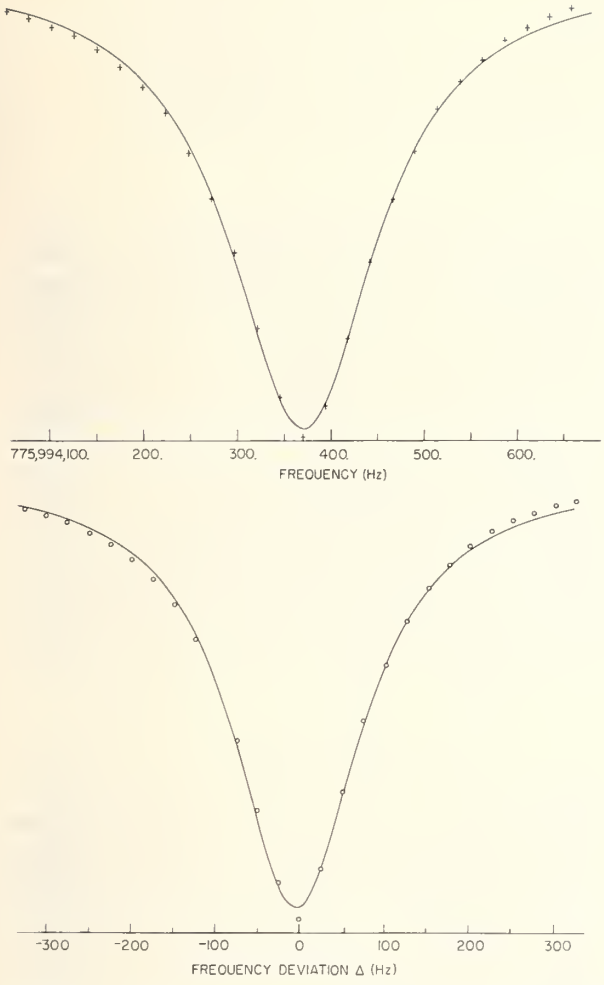


FIGURE 5. *Experimental data points (upper plot) and data points calculated by integrating the equations of motion for an assumed model of the magnetic field gradients (lower plot).* The curves are Lorentzian with a dispersive component.

averaging. The curves are fitted to a Lorentzian with a dispersive component.

Although the computer-generated curves indicated that the pulling effects vanish with decreasing proton line width, they could not show exactly how they vanish. On the basis of various hand waving arguments one can make a good case that the shifts must be proportional to the square of the proton line width, and within the resolution of the computer generated curves; this indeed was the case.

Regardless of the details of the line shape, it can be shown from the symmetry of the governing equations that the two proton transitions are pulled in opposite directions if the experimental conditions are identical. However, in general only one transition was observed during a run due to time limitations. (The work had to be carried out during the wee hours of the morning, for the usual reasons.) In spite of this, one can argue that whatever effect might be left over by the extrapolation procedure should be

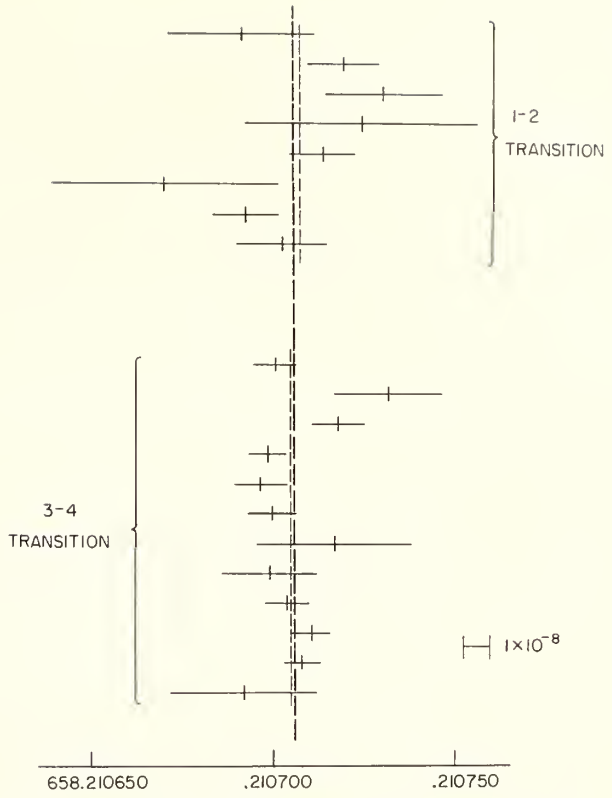


FIGURE 6. *Final results for data fitted to pure Lorentzians and extrapolated quadratically to zero line width.* The upper group is for the transition at ν_{p1}, and the lower group is for the transition ν_{p2}.

treated as a random error. In particular, the shifts are of either sign and vary from day to day. The data shown in figure 6 appears to substantiate that the residual errors are randomly distributed. Here, the data have been fitted by simple Lorentzian curves, and the centers of the curves for each run have been extrapolated quadratically to zero line width. The two proton transitions are grouped separately, and the scatter of data within each group, and between the two groups, appears random. Table Ia gives the numerical results.

As a check on our procedure, we analyzed the data by an alternative method: the data were fit by a Lorentzian curve with a dispersive component. If the curve were truly of this form then the line center would be independent of the line width. This was not quite the case—there was a residual shift which appeared to be linear in the line width. This was found to be true for both the computer modeled data and the real data. Figure 7 shows the results of fitting the data to a dispersive Lorentzian line and extrapolating the line center linearly in the line width. The results are also summarized in table Ib. For reasons which are not clear the standard deviation with this method is about twice as large as with the previous method. If goodness-of-fit were the criteria for selecting an extrapolation procedure then clearly this method would be discarded. However, there

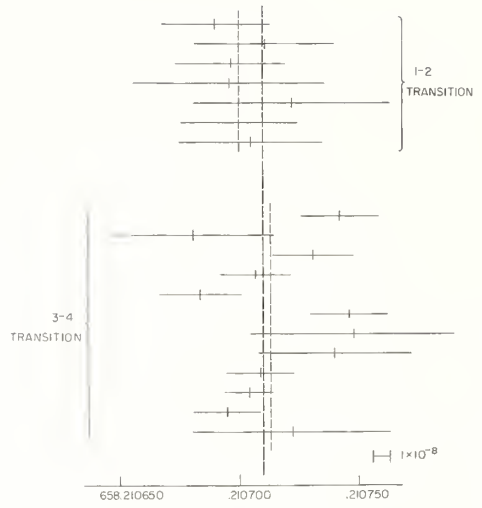


FIGURE 7. *The same data as in figure 6, but fitted to a Lorentzian with a dispersive component, the final results being extrapolated linearly to zero line width.*

seems little justification for such a rash act, particularly since there is a discrepancy of 8×10^{-9} between the two methods. Although the original extrapolation method appears to have better theoretical justification, the discrepancy is cause for unease.

There are two other major sources of systematic error: spin exchange frequency shifts and the wall shift. Both of these shift the two proton transitions in opposite directions, and should average to zero. Tables Ia and Ib show that the results for the two proton transitions are in close agreement, so that there is no evidence for these effects.

A further check on the absence of systematic effects is the consistency between ν_{p1} and ν_{p2}. Their sum should be the hyperfine frequency, $\Delta\nu(\mathrm{H})$. We find

$$\nu_{p1}+\nu_{p2}=1\ 420\ 405\ 751.84(11)\ \mathrm{Hz}$$

in agreement with the present value

$$\Delta\nu(\mathrm{H})=1\ 420\ 405\ 751.769(4)\ \mathrm{Hz}.$$

TABLE 1. *Experimental results for* $\mu_j(\mathrm{H})/\mu_p(\mathrm{H})$

Ia. Results for fitting the data by a Lorentzian curve and extrapolating quadratically to zero line width. Errors are one standard deviation.

Transition	$-\mu_j(\mathrm{H})/\mu_p(\mathrm{H})$	Fractional error
ν_{p1}	658.210 707 5(44)	6.7×10^{-9}
ν_{p2}	658.210 705 0(19)	2.9×10^{-9}
Both	658.210 705 4(18)	2.7×10^{-9}

Ib. Results for fitting the data to a Lorentzian curve with a dispersive component, and extrapolating linearly to zero line width.

ν_{p1}	658.210 700 2(104)	16×10^{-9}
ν_{p2}	658.210 712 7(49)	7.4×10^{-9}
Both	658.210 710 4(44)	6.7×10^{-9}
Final result	658.210 706 1(65)	10×10^{-9}
(Previous result)	658.210 49 (20)	

5. Results

Assigning a final uncertainty to our results presents the usual dilemma associated with precision measurements. Tables Ia and Ib present slightly different, though self-consistent, pictures of the results. Since there seems to be no *a priori* rule for choosing between them, we take their weighted averages as the best estimate of the magnetic moment ratio. The result, 658.210 706 1, is in excellent agreement with the results of table Ia, well within the standard deviation of 2.7×10^{-9}, and differs by exactly one standard deviation (6.7×10^{-9}) from the results of table Ib. However, this may be fortuitous. When all is said and done, we feel that the results are probably reliable to one part in 10^8, but that it is unjustified to press beyond that point because of uncertainties in the extrapolation method. Hence, the final result is

$$-\frac{\mu_j(\mathrm{H})}{\mu_p(\mathrm{H})}=-\frac{g_j(\mathrm{H})}{g_p(\mathrm{H})}=658.210\ 706\ 1(65)$$

where the bracketed quantity is an estimated uncertainty.

To obtain the proton moment in Bohr magnetons from this result, it is necessary to correct for binding effects on the proton and electron, and for the anomalous moment of the electron. All of these are discussed elsewhere in these proceedings, the former by Grotch and Hegstrom, and the latter by Rich and Brodsky. Incidentally, in the interpretive notation described in the foreword, we have

$$-g_j(\mathrm{H})/g_p(\mathrm{H})=6\mathrm{X}58!.21\underline{0}\ \underline{7}!\mathrm{x}\uparrow 06!?\ 1?(65?).$$

6. Acknowledgments

Much of this work was executed by P. Frank Winkler, and is described in a thesis by him [7]. Valuable contributions were also made by Than Myint and Frederick G. Walther. The work was supported by the MIT Sloan Fund for Basic Research, the Office of Naval Research, and the Joint Services Electronics Program.

7. References

[1] Parker, W. H., Taylor, B. N., and Langenberg, D. N., Phys. Rev. Letters **18**, 287 (1967).
[2] Klein, E., Z. Physik **208**, 28 (1968).
[3] Lambe, E. B. D., Polarisation, Matiere, et Rayonnement (Societe Francoise de Physique, Paris), p. 441.
[4] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969).
[5] Myint, T., Kleppner, D., Ramsey, N. F., and Robinson, H. G., Phys. Rev. Letters **17**, 405 (1966).
[6] Brenner, D., Phys. Rev. **185**, 26 (1969).
[7] Winkler, P. F., Thesis, Harvard University, 1970 (unpublished).

$1S_{1/2}$ Bound State Corrections to the Electron and Proton g Factors for Atomic Hydrogen

Roger A. Hegstrom

Department of Chemistry, Wake Forest University, Winston-Salem, North Carolina 27109

Bound state corrections of orders α^2, α^3, $\alpha^2 m/M$ and higher are calculated for the electron and proton g factors in atomic hydrogen. The method of calculation involves (1) finding a set of mutually commuting operators including the Breit Hamiltonian, (2) partially separating the usual center-of-mass coordinates by choosing the wavefunction to be a simultaneous eigenfunction of these operators (3) performing a Foldy-Wouthuysen transformation of the resulting simplified Breit equation, and (4) using first order perturbation theory, with the nonrelativistic Pauli-Schrödinger eigenfunction as the zero-order wave function, to obtain the g factors. Results are also applied to deuterium, and comparisons are made with experiments and with earlier theoretical results.

Key words: Electron; g factor; hydrogen atom; proton; Zeeman effect.

1. Introduction

Recent precision measurements [1–3] of g factor ratios for the $1S_{1/2}$ ground states of atomic hydrogen and deuterium have stimulated new theoretical investigations of higher order atomic bound state corrections to electron and nuclear g values. This author [4] and H. Grotch [5] have calculated hydrogenic $1S_{1/2}$ bound state corrections to the electron g factor to orders α^2, α^3, $\alpha^2 m/M$ and higher, and this author has also calculated corrections to the proton g factor for atomic hydrogen [4] (α is the fine structure constant, m and M are the masses of the electron and proton, respectively). In this paper we employ a slightly different method of calculation which takes into account constants of the motion for the neutral atom in a constant, uniform magnetic field. The results agree with the earlier calculations [4, 5] in orders α^2 and α^3, but modify the earlier mass corrections for both electron and proton g factors in order $\alpha^2 m/M$ and higher. The modification of the $\alpha^2 m/M$ term for the electron g factor confirms a recent calculation by Grotch [6].

In section 2 we find for a neutral atom a set of mutually commuting operators which includes the relativistic Hamiltonian. We then take the exact wavefunction to be a simultaneous eigenfunction of these operators. For this wavefunction the usual center-of-mass (CM) variables are separated, and a new relativistic equation involving only the coordinate $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ results.

In section 3 we perform a Foldy-Wouthuysen transformation on the relativistic equation and calculate bound state corrections to the g factors for the two particles. The results are applied to atomic hydrogen and deuterium, although they also may be applied more generally. Finally, these results are discussed in section 4.

2. Separation of Center of Mass Coordinates for the Breit Equation

Our starting point is the same as that of ref. [4], namely the Breit equation for two particles in a constant, uniform magnetic field, with anomalous magnetic moments introduced phenomenologically as Pauli moments:

$$\mathcal{H}\Psi = E\Psi, \tag{1}$$

where

$$\mathcal{H} = \boldsymbol{\alpha}_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)+\boldsymbol{\alpha}_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2) + \beta_1 m_1+\beta_2 m_2+B(\mathbf{r}), \tag{2}$$

and $B(\mathbf{r})$ is the Breit operator plus anomalous moment terms:

$$\begin{aligned} B(\mathbf{r}) = {} & (e_1e_2/r)\,(1-\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2/2-\boldsymbol{\alpha}_1\cdot\mathbf{r}\boldsymbol{\alpha}_2\cdot\mathbf{r}/2r^2) \\ & -a_1(e_1/2m_1)\left[\beta_1\boldsymbol{\sigma}_1\cdot(\mathbf{H}+e_2\boldsymbol{\alpha}_2\times\mathbf{r}/r^3)-i\beta_1e_2\boldsymbol{\alpha}_1\cdot\mathbf{r}/r^3\right] \\ & -a_2(e_2/2m_2)\left[\beta_2\boldsymbol{\sigma}_2\cdot(\mathbf{H}-e_1\boldsymbol{\alpha}_1\times\mathbf{r}/r^3)+i\beta_2e_1\boldsymbol{\alpha}_2\cdot\mathbf{r}/r^3\right]. \end{aligned} \tag{3}$$

Quantities appearing in eqs (1) and (2) have been defined previously [4]. Here we use natural units and choose the gauge:

$$\mathbf{A}_i=\tfrac{1}{2}\mathbf{H}\times\mathbf{r}_i, \qquad i=1,2. \tag{4}$$

We now define an operator

$$\tilde{\mathbf{\Pi}}=\mathbf{p}_1+e_1\mathbf{A}_1+\mathbf{p}_2+e_2\mathbf{A}_2. \tag{5}$$

It can be easily shown that components of $\tilde{\Pi}$ satisfy the commutation relations

$$[\tilde{\Pi}_i, \mathcal{H}]=0,$$
$$[\tilde{\Pi}_i, \tilde{\Pi}_j]=-i(e_1+e_2)\epsilon_{ijk}H_k, \tag{6}$$

where ϵ_{ijk} is the Levi-Civita symbol. For $e_1+e_2=0$, we can choose Ψ to be a simultaneous eigenfunction of $\mathcal{H}$ and $\tilde{\Pi}$. Such a wave function will have the form [7]

$$\Psi=\exp(i\mathbf{\Pi}\cdot\mathbf{R})\psi(\mathbf{r}) \tag{7}$$

where

$$\mathbf{\Pi}\equiv\mathbf{K}-e_1\mathbf{A}_1-e_2\mathbf{A}_2=\mathbf{K}-\tfrac{1}{2}e_1\mathbf{H}\times\mathbf{r}. \tag{8}$$

K is the eigenvalue of $\tilde{\Pi}$:

$$\tilde{\mathbf{\Pi}}\Psi=\mathbf{K}\Psi, \tag{9}$$

R is the usual CM variable

$$\mathbf{R}=(m_1\mathbf{r}_1+m_2\mathbf{r}_2)/(m_1+m_2) \tag{10}$$

and $\psi(\mathbf{r})$ is a function of $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$ only. By allowing the Breit Hamiltonian [eq (2)] to operate on Ψ [eq (7)], one can show that if Ψ is to satisfy the Breit equation (1) for the eigenvalue E, ψ must satisfy the following relativistic equation:

$$[\boldsymbol{\alpha}_1\cdot\boldsymbol{\pi}_1+\boldsymbol{\alpha}_2\cdot\boldsymbol{\pi}_2+\beta_1m_1+\beta_2m_2+B(\mathbf{r})]\psi(\mathbf{r})=E\psi(\mathbf{r}). \tag{11}$$

Here

$$\boldsymbol{\pi}_1\equiv m_1\mathbf{K}/(m_1+m_2)+\mathbf{p}-e_1\mathbf{A}(\mathbf{r})$$
$$\boldsymbol{\pi}_2\equiv m_2\mathbf{K}/(m_1+m_2)-\mathbf{p}-e_2\mathbf{A}(-\mathbf{r}), \tag{12}$$

p is the relative momentum

$$\mathbf{p}=(m_2\mathbf{p}_1-m_1\mathbf{p}_2)/(m_1+m_2), \tag{13}$$

and

$$\mathbf{A}(\mathbf{r})=\tfrac{1}{2}\mathbf{H}\times\mathbf{r}. \tag{14}$$

We have continued to distinguish the electric charges e_1 and e_2, although eqs (7)–(10) and what follows hold only for $e_1+e_2=0$.

Solving the relativistic equation (11) is equivalent to solving the Breit equation (1). It is interesting to note that eq (11) is similar in form to the Breit equation except for the definitions (12) of $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ and for the fact that eq (11) involves only the coordinate $\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2$.

3. Reduction of the Relativistic Equation and Calculation of Bound State Corrections to the g Factors

We now proceed to perform a reduction of the relativistic eq (11) by means of the Foldy-Wouthuysen method. This is the same procedure employed in ref. [4] and follows closely the treatment of Barker and Glover [8]. We find for the transformed Hamiltonian

$$\mathcal{H}'=m_1+m_2+\sum_{i=0}^{6}\mathcal{H}_i, \tag{15}$$

where

$$\mathcal{H}_0=\boldsymbol{\pi}_1^2/2m_1+\boldsymbol{\pi}_2^2/2m_2+e_1e_2/r$$
$$\mathcal{H}_1=-\boldsymbol{\pi}_1^4/8m_1^3-\boldsymbol{\pi}_2^4/8m_2^3$$
$$\mathcal{H}_2=-\frac{\pi e_1e_2}{2m_1^2}(1+2a_1)\delta(\mathbf{r})-\frac{\pi e_1e_2}{2m_2^2}(1+2a_2)\delta(\mathbf{r})$$
$$\mathcal{H}_3=-\frac{e_1e_2}{4m_1^2}(1+2a_1)\boldsymbol{\sigma}_1\cdot\frac{\mathbf{r}\times\boldsymbol{\pi}_1}{r^3}+\frac{e_1e_2}{4m_2^2}(1+2a_2)$$
$$\times\boldsymbol{\sigma}_2\cdot\frac{\mathbf{r}\times\boldsymbol{\pi}_2}{r^3}+\frac{e_1e_2}{2m_1m_2}(1+a_1)\boldsymbol{\sigma}_1\cdot\frac{\mathbf{r}\times\boldsymbol{\pi}_2}{r^3}$$
$$-\frac{e_1e_2}{2m_1m_2}(1+a_2)\boldsymbol{\sigma}_2\cdot\frac{\mathbf{r}\times\boldsymbol{\pi}_1}{r^3}$$
$$\mathcal{H}_4=-\frac{e_1}{2m_1}\left\{\boldsymbol{\sigma}_1\cdot\mathbf{H}\left(1-\frac{\boldsymbol{\pi}_1^2}{2m_1^2}\right)+a_1\left(\boldsymbol{\sigma}_1\cdot\mathbf{H}-\frac{\boldsymbol{\sigma}_1\cdot\boldsymbol{\pi}_1\boldsymbol{\pi}_1\cdot H}{2m_1^2}\right)\right\}$$
$$-\frac{e_2}{2m_2}\left\{\boldsymbol{\sigma}_2\cdot\mathbf{H}\left(1-\frac{\boldsymbol{\pi}_2^2}{2m_2^2}\right)+a_2\left(\boldsymbol{\sigma}_2\cdot\mathbf{H}-\frac{\boldsymbol{\sigma}_2\cdot\boldsymbol{\pi}_2\boldsymbol{\pi}_2\cdot\mathbf{H}}{2m_2^2}\right)\right\}$$
$$\mathcal{H}_5=-\frac{e_1e_2}{4m_1m_2}(1+a_1)(1+a_2)$$
$$\times\left\{\tfrac{8}{3}\pi\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\delta(r)+\frac{3(\boldsymbol{\sigma}_1\cdot\mathbf{r})(\boldsymbol{\sigma}_2\cdot\mathbf{r})-\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2r^2}{r^5}\right\}$$
$$\mathcal{H}_6=-(e_1e_2/2m_1m_2)\{r^{-1}\boldsymbol{\pi}_1\cdot\boldsymbol{\pi}_2+\mathbf{r}[(\mathbf{r}/r^3)\cdot\boldsymbol{\pi}_1]\cdot\boldsymbol{\pi}_2\} \tag{16}$$

The transformed Hamiltonian also contains terms of orders $m\alpha^6$, $eH\alpha^4/m$ and higher which we have not included in eq (16).

To calculate bound state corrections to the g factors, we procced as in ref. [4], using first order perturbation theory with the nonrelativistic Pauli-Schrödinger eigenfunctions for the $1S_{1/2}$ ground state as the unperturbed functions. Setting the eigenvalue K equal to zero, we evaluate the contributing terms:

$$\langle 1S\mid\mathcal{H}_3+\mathcal{H}_4\mid 1S\rangle$$
$$=-\left(\frac{e_1}{2m_1}\right)\tfrac{1}{2}g_1\langle\boldsymbol{\sigma}_1\cdot H\rangle\left\{1-\tfrac{1}{2}\alpha^2\left(\frac{m_2}{m_1+m_2}\right)^2\left(\frac{1+\frac{1}{3}a_1}{1+a_1}\right)\right.$$
$$\left.+\tfrac{1}{6}\alpha^2\left(\frac{m_2}{m_1+m_2}\right)\left(\frac{1+2a_1}{1+a_1}\right)-\tfrac{1}{3}\alpha^2\left(\frac{m_1}{m_1+m_2}\right)\right\}$$
$$-\left(\frac{e_2}{2m_2}\right)\tfrac{1}{2}g_2\langle\boldsymbol{\sigma}_2\cdot\mathbf{H}\rangle\left\{1-\tfrac{1}{2}\alpha^2\left(\frac{m_1}{m_1+m_2}\right)^2\left(\frac{1+\frac{1}{3}a_2}{1+a_2}\right)\right.$$
$$\left.+\tfrac{1}{6}\alpha^2\left(\frac{m_1}{m_1+m_2}\right)\left(\frac{1+2a_2}{1+a_2}\right)-\tfrac{1}{3}\alpha^2\left(\frac{m_2}{m_1+m_2}\right)\right\} \tag{17}$$

where $g_i=2(1+a_i)$ is the g factor for the free particle $i=1, 2$. The g factor for the bound particle i is then given by g_i times the appropriate expression in curly brackets in eq (17). For atomic hydrogen, with $m_1=m$, $m_2=M$, $a_e=\alpha/2\pi-0.328\alpha^2/\pi^2$ and $a_p=1.793$,

expanding to orders α^3 and $\alpha^2 m/M$ we find

$$g_e(1S) = g_e\left\{1-\tfrac{1}{3}\alpha^2+\frac{1}{4\pi}\alpha^3+\tfrac{1}{2}\alpha^2\frac{m}{M}\right\}$$

$$g_p(1S) = g_p\left\{1-\tfrac{1}{3}\alpha^2+\tfrac{1}{2}\alpha^2\frac{m}{M}\left(\frac{1+\frac{4}{3}a_p}{1+a_p}\right)\right\}. \quad (18)$$

From eq (18) we may calculate expressions for the g factor ratios corresponding to the experimental measurements for atomic hydrogen and deuterium [1–3]:

$$\frac{g_e(1S, H)}{g_p(1S, H)} = \frac{g_e}{g_p}\left\{1+\frac{1}{4\pi}\alpha^3-\tfrac{1}{6}\alpha^2\frac{m}{M}\left(\frac{a_p}{1+a_p}\right)\right\}$$
$$= (g_e/g_p)(1+0.28\cdot 10^{-7}) \quad (19)$$

$$g_e(1S, H)/g_e(1S, D) = 1+\tfrac{1}{4}\alpha^2(m/M)$$
$$= 1+7.25\cdot 10^{-9}. \quad (20)$$

We conclude this section by stating that it is also possible to derive energy expressions for the hyperfine levels of the atom in a magnetic field by evaluating matrix elements for the other perturbations appearing in eq (16) and diagonalizing the resulting Hamiltonian matrix in the usual way. To first order, the result is the usual Breit-Rabi formula [9]

$$W(F, m_F) = W_0-\tfrac{1}{4}\nu_0-g_p(1S)\mu_n H m_F \pm\tfrac{1}{2}\nu_0(1+2m_F x+x^2)^{1/2}, \quad (21)$$

where

$$W_0 = \langle 1S \mid \mathcal{H}_0+\mathcal{H}_1+\mathcal{H}_2 \mid 1S\rangle \quad (22)$$

is a term which shifts each level by the same amount, where

$$\nu_0 = 4\langle 1S, m_F=1 \mid \mathcal{H}_5 \mid 1S, m_F=1\rangle$$
$$= (2/3)g_e g_p \alpha^4 m^2/M \quad (23)$$

is the Fermi hyperfine energy, and where

$$x = [g_p(1S)\mu_n+g_e(1S)\mu_0]H/\nu_0. \quad (24)$$

The g factors appearing above are defined by eq (18). Thus we have explicitly included higher order corrections to the g factors in the Breit-Rabi formula. It is not apparent from the methods of this paper, however, that the energy will continue to have the Breit-Rabi form in yet higher orders. Indeed, some of the operators in eq (16) mix in excited electronic states in higher orders of perturbation theory. Since the Breit-Rabi formula is often used to interpret the results of high-precision experiments [10], this question seems worth future investigation.

4. Discussion

The results of this paper, eq (17), are in the same form as those of ref. [4]. The last three terms in each pair of curly brackets may be interpreted as arising from a relativistic mass correction, a spin-orbit coupling term, and a spin-other-orbit term, respectively. These results also differ in one way from those of ref. [4], however: in eq (17) the mass factors $m_1/(m_1+m_2)$ and $m_2/(m_1+m_2)$ in the spin-orbit and spin-other-orbit terms are not squared as they are in ref. [4]. This difference arises basically because the results of ref. [4] apply to the CM frame with $\mathbf{P}\equiv\mathbf{p}_1+\mathbf{p}_2=0$. To orders α^2 and α^3, this is a good approximation. It is generally not valid, however, to set $\mathbf{P}=0$ since $\mathbf{P}$ is not a constant of motion in the magnetic field. In the present treatment setting $\mathbf{K}=0$ is valid because the operator $\tilde{\mathbf{\Pi}}$ is a constant of motion [11].

Now, eq (19) may be used to correct the experimentally determined [1] electron-proton g factor ratio for atomic hydrogen to obtain an accurate value for the free-particle ratio g_e/g_p. The experiment [1] does not test the results of these calculations. On the other hand, the measurements [2, 3] of the ratio of electron g values in hydrogen and deuterium do provide a direct experimental test of these results. Our theoretical result $g_e(1S, H)/g_e(1S, D)-1 = 7.25\cdot 10^{-9}$ may be compared with the experimental results $(7.2\pm 1.2)\cdot 10^{-9}$ [2] and $(9.4\pm 1.4)\cdot 10^{-9}$ [3].

In conclusion, we note that an interesting result of this treatment is a simple physical interpretation for the bound state corrections to the g factors of the two particles. This interpretation remains essentially unaltered from that presented in ref. [4], where the corrections are interpreted as arising from relativistic mass, spin-orbit, and spin-other-orbit effects. This interpretation also provides an explanation for the somewhat surprising equality of the electron and proton corrections in atomic hydrogen to order α^2 [1]. To this order, the correction to the electron g factor receives a contribution $-\frac{1}{2}\alpha^2$ due to relativistic mass change and a contribution $+\alpha^2/6$ due to spin-orbit coupling; the correction to the proton g factor, on the other hand, arises entirely from the spin-other-orbit contribution $-\alpha^2/3$ (also called the "Lamb diamagnetic shielding" correction). In higher order, all three effects contribute to both g factors and the corrections are no longer equal.

5. Acknowledgment

The author wishes to acknowledge helpful correspondence with Professor H. Grotch.

6. References

[1] Myint, T., Kleppner, D., Ramsey, N. F., and Robinson, H. G., Phys. Rev. Letters **17**, 405 (1966). See also D. Kleppner, these Proceedings.
[2] Hughes, W. M., and Robinson, H. G., Phys. Rev. Letters **23**, 1209 (1969). See also H. G. Robinson and W. M. Hughes, these Proceedings.
[3] Larson, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23**, 1369 (1969).
[4] Hegstrom, R. A., Phys. Rev. **184**, 17 (1969) and corrections in Phys. Rev. **A1**, 536 (1970).
[5] Grotch, H., Phys. Rev. Letters **24**, 39 (1970).
[6] Grotch, H., private communication and submitted to Phys. Rev. Comments. See also H. Grotch, these Proceedings.

[7] A similar result has been obtained in the nonrelativistic case by Lamb, W. E., Jr., Phys. Rev. **85**, 259 (1952), and by Carter, B. P., J. Math. Phys. **10**, 788 (1969).
[8] Barker, W. A., and Glover, F. N., Phys. Rev. **99**, 317 (1955). Note that an additional anomalous moment term must appear in the Breit Hamiltonian [see Barker and Glover's eq (16a)] to obtain the complete hyperfine term H_5.
[9] See for example, Ramsey, N. F., Molecular Beams (Oxford University Press, London, 1956), p. 80.
[10] See for example, Myint, T., Thesis, Harvard Univ. (1966).
[11] Setting $K=0$ is not necessary, but is done here for convenience. This represents a "non-moving" atom in the lab frame. With $K\neq 0$, we obtain the usual motional Stark effect (see ref. [7]) and other motional terms.

DISCUSSION

S. J. Brodsky: I might mention that when you use the Foldy-Wouthysen Hamiltonian for a two-body system in an external field that you have to be careful only in the case of an external electric field. It just breaks down and you have to be much more careful. However, the computations for an external magnetic field are usually all right.

Electromagnetic Interactions of Hydrogenic Atoms: Corrections to g Factors

Howard Grotch

Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802

Corrections to g factors in hydrogenic atoms are obtained from the Breit equation. Explicit wave functions for a slowly moving atom are obtained in the absence of external fields. Interaction with a constant magnetic field is treated perturbatively to obtain bound state corrections to g factors arising from relativistic effects and from nuclear motion. Another method involving a unitary transformation is also discussed. The advantage of this approach is that for a neutral system the unperturbed wave functions are given in the rest frame. Results are compared with experiment; in particular the ratio of $g_e(H)/g_e(D)$.

Key words: Bound state corrections; g factors in hydrogenic atoms; nuclear motion; ratio of $g_e(H)/g_e(D)$; relativistic effects.

1. Introduction

The Breit equation has been used extensively in atomic physics to treat problems in which relativistic effects are important. It is well-known that it lacks the manifest covariance which is present in the Bethe-Salpeter equation, and therefore it can only be regarded as an approximate equation. Nevertheless it does provide a good starting point for the solution of atomic problems. In fact, if we consider hydrogen as an example, we find that the Breit equation is in a sense a better starting point than is the Bethe-Salpeter equation in ladder approximation, for whereas the former reduces to the Dirac equation for the electron (in the static limit) the latter does not.

The Breit Hamiltonian for an atom in an external electromagnetic field may be obtained through the usual minimal coupling prescription of the Breit Hamiltonian in the absence of such a field. If the external field is weak, its effects may be calculated perturbatively. This requires a knowledge of the unperturbed wave function. Since the external field terms can in general have matrix elements between atomic states of different momenta, it is necessary to obtain unperturbed wave functions which describe a moving atom. In the nonrelativistic domain, the separability of the Schrödinger equation readily enables us to construct the wave function for a moving atom once we know the relative wave function. In the relativistic domain, however, separability is not possible and therefore knowledge of the relative wave function does not automatically yield the whole wave function.

In this work we show that it is possible to obtain a solution of the Breit equation which describes a moving atom. For simplicity we restrict our discussion to the hydrogen atom. After obtaining suitable unperturbed wave functions we calculate the change in the energy levels of the atom for the case of a constant magnetic field. We obtain the electron g factor in the ground state of hydrogen, including binding corrections and nuclear mass corrections. In previous work [1, 2] the mass corrections were not properly evaluated, and therefore the present paper gives new results for these corrections.[1] There are also perturbations arising from radiative corrections. These also give rise to Zeeman splittings as discussed in ref. [1]. However, again the nuclear mass dependence of these corrections was not properly obtained and the new result is given here.

The calculation, done in the manner suggested above, requires a knowledge of the wave function for a moving atom. However, we have found that another method may be used which greatly simplifies some of the technical aspects of the calculation. In effect we make a unitary transformation of the Hamiltonian so that the transformed Hamiltonian does not contain the center of mass position operator. Having done this the perturbation theory is obviously nondegenerate between eigenstates of the operator $\mathbf{P}$; moreover we may then use wave functions for an atom at rest. Hegstrom [3] has also discussed this method.

We summarize the results by giving the electron g factor. Since the proton g factor has also been calculated by Hegstrom [3] and by this author using the Barker-Glover [4] method, we present the results of that calculation as well. Finally we compare the numerical results of $g_e(H)/g_e(D)$ obtained theoretically with recent measurements of this ratio.

[1] Corrected results have also been given in H. Grotch, Phys. Rev. A2, 1605 (1970). See Ref. [7].

2. Approximate Solution of the Breit Equation

The Breit Hamiltonian for the hydrogen atom is

$$H_0=\alpha_e\cdot\mathbf{p}_e+\beta_e m+\alpha_p\cdot\mathbf{p}_p+\beta_p M-\frac{Z\alpha}{r}$$
$$+\left(\frac{Z\alpha}{2r}\right)\left(\alpha_e\cdot\alpha_p+\frac{\alpha_e\cdot\mathbf{r}\alpha_p\cdot\mathbf{r}}{r^2}\right). \quad (1)$$

Z is unity in hydrogen but for the purpose of distinguishing between powers of α which arise from radiative corrections and powers of α which arise from binding, we use $Z\alpha$ to denote the binding terms.

Since the proton mass M is much larger than the Bohr momentum, the nucleus may be treated nonrelativistically. If we restrict ourselves to the large component space of the proton and if we further ignore all spin-dependent terms of the proton we obtain from eq (1)

$$H=\alpha_e\cdot\mathbf{p}_e+\beta_e m+M+\frac{\mathbf{p}_p^2}{2M}-\frac{Z\alpha}{r}$$
$$+\frac{Z\alpha}{2Mr}\left(\alpha_e\cdot\mathbf{p}_p+\frac{1}{r^2}\alpha_e\cdot\mathbf{r}\mathbf{r}\cdot\mathbf{p}_p\right). \quad (2)$$

We may, of course, treat the proton spin-dependent terms as perturbations and may also recover the lower components of the proton wave function when needed.

H may be separated into relative parts and parts which involve $\mathbf{P}$. We use the separation

$$\mathbf{p}_e=\mathbf{p}+\frac{m}{m+M}\mathbf{P} \quad \text{and} \quad \mathbf{p}_p=-\mathbf{p}+\frac{M}{m+M}\mathbf{P}. \quad (3)$$

Therefore

$$H=\alpha_e\cdot\mathbf{p}+\beta_e m+V+M+\frac{\mathbf{p}^2}{2M}$$
$$+\frac{V}{2M}\left[\alpha_e\cdot\mathbf{p}+\frac{1}{r^2}\alpha_e\cdot\mathbf{r}\mathbf{r}\cdot\mathbf{p}\right]$$
$$+\frac{m}{m+M}\alpha_e\cdot\mathbf{P}-\frac{\mathbf{p}\cdot\mathbf{P}}{m+M}+\left(\frac{M}{m+M}\right)\frac{\mathbf{P}^2}{2(m+M)}$$
$$-\frac{V}{2(m+M)}\left[\alpha_e\cdot\mathbf{P}+\frac{1}{r^2}\alpha_e\cdot\mathbf{r}\mathbf{r}\cdot\mathbf{P}\right]$$
$$\equiv H_{\rm rel}+H_P, \quad (4)$$

where H_P contains only $\mathbf{P}$ dependent terms and $H_{\rm rel}$ contains the relative part; V is $-Z\alpha/r$.

Approximate eigenfunctions of $H_{\rm rel}$ have been given by Grotch and Yennie [5]. It should be noted that in obtaining these solutions $H_{\rm rel}$ was rearranged and certain terms proportional to $1/M^2$ were discarded. Therefore in doing calculations with these wave functions we must bear in mind that certain corrections involving $1/M^2$ will be missing. We will discuss this in more detail later. Using the wave functions of ref. [5] we may construct the wave function for the moving atom. It may readily be shown that if

$$H_{\rm rel}\psi_0(\mathbf{r})=\mathfrak{M}\psi_0(\mathbf{r}) \quad (5)$$

then

$$H\psi(\mathbf{r},\mathbf{X})=\{\mathfrak{M}+[\mathbf{k}^2/2(m+M)]\}\psi(\mathbf{r},\mathbf{X}) \quad (6)$$

if

$$\psi(\mathbf{r},\mathbf{X})=[1+(\alpha\cdot\mathbf{k}/2\mathfrak{M})+(i\eta\mathbf{k}\cdot\mathbf{r}/\mathfrak{M})]$$
$$\times\psi_0(\mathbf{r})\exp(i\mathbf{k}\cdot\mathbf{X}), \quad (7)$$

where $\eta=\epsilon-V/2$, ϵ being the basic nonrelativistic binding energy associated with the relative wave function

$$\psi_0(\mathbf{r}); \quad \text{i.e. } \epsilon\simeq\mathfrak{M}-m-M.$$

In proving that eq (7) is a solution of eq (6) we have ignored terms of order $mk\epsilon/M^2$ and $k^2\epsilon/M^2$. Therefore in eq (6) the term $m+M$ could just as well be $\mathfrak{M}$.

The wave function of eq (7) describes an unperturbed hydrogen atom quite well and may be used to do perturbation theory. In the next section we discuss the example of interaction with a constant magnetic field.

3. Perturbation Theory in a Constant Magnetic Field

To introduce a constant magnetic field we replace $\mathbf{p}_e$ and $\mathbf{p}_p$ by $\mathbf{p}_e-e\mathbf{A}_e$ and $\mathbf{p}_p+Ze\mathbf{A}_p$ respectively, where $\mathbf{A}_{e,p}$ is $\frac{1}{2}\mathbf{H}\times\mathbf{x}_{e,p}$. From eq (2) we see that there is a magnetic interaction linear in $\mathbf{H}$ of

$$H_{\rm mag}=-e\alpha_e\cdot\mathbf{A}_e+(Ze/M)\mathbf{p}_p\cdot\mathbf{A}_p+(Z^2\alpha e/2Mr)$$
$$\times[\alpha_e\cdot\mathbf{A}_p+(1/r^2)\alpha_e\cdot\mathbf{r}\mathbf{r}\cdot\mathbf{A}_p]. \quad (8)$$

To obtain the change in the energy levels caused by $H_{\rm mag}$ we use first-order nondegenerate perturbation theory between the wave functions of eq (7). It may be shown that this is a valid procedure since in first-order the interaction due to a constant magnetic field does not connect states of different center of mass momenta.

In previous work [2] we used wave functions given by Brodsky and Primack [6]. We subsequently learned that these wave functions were not accurate enough to yield all corrections of relative order $(Z\alpha)^2m/M$ to the electron g factor. Therefore the nuclear mass corrections given in ref. [2] require modification. We expect that the wave function given above in eq (7) is accurate enough to yield corrections of the above magnitude, but that it is

not accurate enough to yield terms of relative order $(Z\alpha)^2m^2/M^2$. This is due to the fact that certain terms in $H_{\rm rel}$ were discarded in ref. [5], and therefore, as mentioned above, $\psi_0(\mathbf{r})$ is not the exact solution of eq (5).

Each of the terms in eq (8) gives a Zeeman contribution. However, as mentioned earlier, by doing first-order perturbation theory with the wave functions of eq (7) we do not expect to obtain all $1/M^2$ corrections. To be specific relative correction terms of order $(m/M)^2(Z\alpha)^2$ will be missing in the calculation of the electron g factor. Therefore, at this stage we may ignore terms in eq (8) which would also give rise to corrections of this order. The only term which survives in eq (8) is the $-e\boldsymbol{\alpha}_e\cdot\mathbf{A}_e$ term. We have calculated its expectation value in an atomic state in which the center of mass momentum is zero. However in performing the calculation we only let $\mathbf{k}$ approach zero after the calculation is done. If we had set $\mathbf{k}$ equal to zero immediately certain terms in eq (7) would be missing; yet these terms give a contribution even in the limit as $\mathbf{k}$ goes to zero. This is due to the fact that the perturbation involves $\mathbf{X}$ and therefore does not commute with the center of mass momentum.

We will not give the details of the calculation here since they are fairly involved. We hope to present details elsewhere. The result for the change in the energy levels in the presence of a magnetic field is given by [7]

$$\Delta E_e=-(e/2m)\langle\boldsymbol{\sigma}_e\cdot\mathbf{H}\rangle[1-\tfrac{1}{3}(Z\alpha)^2+\tfrac{1}{2}(Z\alpha)^2(m/M)]. \tag{9}$$

We have confirmed this result using a number of different methods which include the Barker-Glover method [4], the Foldy-Wouthuysen transformation [8] and the reduction of the Breit equation to large components.

Terms of order $(Z\alpha)^2(m/M)^2$ may be obtained by including the other terms in eq (8) and by incorporating terms dropped in ref. [5]. However, it appears to be quite straightforward to use either the Barker-Glover [4] or the Foldy-Wouthuysen [8] method since these involve expansions in the inverse masses. Results will be given in the next section.

Regardless of which method of attack is used we have found that the problem is simplified considerably by making a unitary transformation of eq (1) such as was made by Lamb [9]. Let us suppose that the center of mass position operator $\mathbf{X}$ is not a dynamical variable. Then it would be quite useful to make a gauge transformation so that the new vector potential vanishes at the center of mass. The virtue of this is that for a neutral system like hydrogen the center of mass operator does not appear in the new Hamiltonian. Now since $\mathbf{X}$ is in fact a dynamical variable we can not readily perform a gauge transformation as discussed above. However we may utilize a unitary transformation identical in form to the gauge transformation.

If we let

$$\psi=\exp\{i\tfrac{1}{2}e[(M+Zm)/(m+M)]\mathbf{r}\cdot\mathbf{H}\times\mathbf{X}\}\phi \tag{10}$$

then we find that eq (1) with $\mathbf{p}_e$ replaced by $\mathbf{p}_e-e\mathbf{A}_e$ and $\mathbf{p}_p$ by $\mathbf{p}_p+Ze\mathbf{A}_p$ transforms to

$$\begin{aligned}H'=&\boldsymbol{\alpha}_e\cdot\left[\mathbf{p}_e-e\mathbf{A}+\frac{m(Z-1)}{m+M}\,e\mathbf{A}_p\right]+\beta_e m\\&+\boldsymbol{\alpha}_p\cdot\left[\mathbf{p}_p-Ze\mathbf{A}+\frac{M}{m+M}\,(Z-1)e\mathbf{A}_e\right]+\beta_p M\\&-\frac{Z\alpha}{r}+\left(\frac{Z\alpha}{2r}\right)\left[\boldsymbol{\alpha}_e\cdot\boldsymbol{\alpha}_p+\frac{\boldsymbol{\alpha}_e\cdot\mathbf{r}\boldsymbol{\alpha}_p\cdot\mathbf{r}}{r^2}\right],\end{aligned} \tag{11}$$

where $\mathbf{A}=\frac{1}{2}\mathbf{H}\times\mathbf{r}$ and $\mathbf{A}_e$ and $\mathbf{A}_p$ have previously been defined. It is obvious from eq (11) that when Z is unity H' does not depend on $\mathbf{X}$, H' commutes with $\mathbf{P}$, and that the use of perturbation theory is greatly simplified. In fact the use of nondegenerate perturbation theory is obvious, and moreover if the atom is at rest only the relative wave function enters into the calculation. In other words we may set $\mathbf{k}$ equal to zero in eq (7) at the outset.

It is interesting to note that before making the unitary transformation, the operator

$$\boldsymbol{\pi}=\mathbf{p}_e+e\mathbf{A}_e+\mathbf{p}_p-Ze\mathbf{A}_p,\qquad Z=1 \tag{12}$$

commuted with the complete Hamiltonian. Under the transformation it went into $\mathbf{P}$ which commutes with the transformed Hamiltonian. The construction of a conserved operator as in eq (12) is not unfamiliar in the nonrelativistic domain. If two noninteracting particles are in a constant magnetic field they each move in Landau orbits and the operators $\mathbf{p}_e+e\mathbf{A}_e$ and $\mathbf{p}_p-Ze\mathbf{A}_p$ are each separately conserved. The transverse components $(p_e+eA_e)_{x,y}$ are proportional respectively to the x and y components of the position vector which locates the center of the electron orbit. The x and y components of this operator do not commute. Similar statements hold for the proton variables. The sum of these operators $\boldsymbol{\pi}$ is also conserved and moreover for Z equal to unity it has commuting components. It is easy to show that even in the relativistic case in which the two particle interaction is specified in eq (1), the operator $\boldsymbol{\pi}$ continues to commute with H if the system is neutral and it continues to have commuting components. Therefore we are doing perturbation theory between eigenstates of $\boldsymbol{\pi}$ and the unitary transformation simplifies the technical aspects of the problem.

As discussed in ref. [2] there are radiative corrections which affect the energy levels of hydrogen in a magnetic field. The contribution from vacuum polarization is zero to the accuracy we require, but the lowest order modification of the electron vertex leads to a correction

$$\Delta E_e^{\rm rad}=\langle(-e/2m)(\alpha/2\pi)(\beta_e\boldsymbol{\sigma}_e\cdot\mathbf{H}_e-i\boldsymbol{\gamma}_e\cdot\mathbf{E}_e)\rangle, \tag{13}$$

where $\mathbf{H}_e$ and $\mathbf{E}_e$ are the magnetic and electric fields at the electron,

$$\mathbf{H}_e=\mathbf{H}-(Ze/4\pi)(\boldsymbol{\alpha}_p\times\mathbf{r}/r^3),$$
$$\mathbf{E}_e=-(Ze/4\pi)(\mathbf{r}/r^3). \tag{14}$$

Some of the terms in eq (13) are not obviously dependent upon $\mathbf{H}$. However it must be remembered that in doing this calculation it is necessary to include the modification of the wave function due to magnetic field terms present in our original Breit equation. These will enter in an important way in the lower components of the electron and proton wave functions. We obtain

$$\Delta E_e^{\rm rad}=\frac{-e}{2m}\frac{\alpha}{2\pi}\langle\boldsymbol{\sigma}_e\cdot\mathbf{H}\rangle\left[1-\tfrac{1}{6}(Z\alpha)^2\left(1-\frac{2m}{M}\right)-\tfrac{1}{3}(Z\alpha)^2\frac{m}{M}+\tfrac{1}{3}(Z\alpha)^2\left(1-\frac{m}{M}\right)\right]$$
$$=\frac{-e}{2m}\frac{\alpha}{2\pi}\langle\boldsymbol{\sigma}_e\cdot\mathbf{H}\rangle\left[1+\tfrac{1}{6}(Z\alpha)^2-\tfrac{1}{3}(Z\alpha)^2\frac{m}{M}\right], \tag{15}$$

There are of course higher order radiative corrections. Since the binding corrections to these will probably bring in additional powers of $(Z\alpha)^2$, then if we ignore terms of order $\alpha^2(Z\alpha)^2$ and smaller we may replace $\alpha/2\pi$ in eq (15) by the complete anomalous moment of the electron.

4. Summary

As a check of the calculations done utilizing the perturbation theory discussed in preceding sections, we have also applied the Barker-Glover method [4] to obtain the change in the energy levels in a constant magnetic field. This is the method which Hegstrom has used [1, 3] to calculate the electron and proton g factors.

In this approach a unitary transformation of the Hamiltonian is made which leads to a transformed Hamiltonian expressed as a sum of terms involving inverse powers of the masses. All terms consist of even operators, which means that they do not connect the upper and lower components of the electron and proton spinors. This particular unitary transformation is made after first applying the unitary transformation discussed in the last section.

Since we have retained terms involving inverse cubes of the masses (e.g., $1/mM^2$, $1/M^3$, etc.), the resulting expressions are given to greater accuracy than eqs (9) and (15). Our results are presented only for a neutral system ($Z=1$). By examining eq (11) it is apparent that additional terms will be present when the system has nonzero total charge. For example we could have a spin $\frac{1}{2}$ nucleus of charge Z. The formulas given below are approximately correct when $Z\neq1$ but there are additional terms which must be added; these are of relative order

$$(Z-1)(Z\alpha)^2m^2/M^2.$$

We obtain for the electron and nuclear g factors in the 1S state

$$g_e(1S)=g_e\left[1-\tfrac{1}{2}(Z\alpha)^2\left(\frac{M}{m+M}\right)^2\left(\frac{1+\frac{1}{3}a_e}{1+a_e}\right)+\tfrac{1}{6}(Z\alpha)^2\times\left(\frac{M}{m+M}\right)\left(\frac{1+2a_e}{1+a_e}\right)-\tfrac{1}{3}(Z\alpha)^2\left(\frac{m}{m+M}\right)\right], \tag{16}$$

$$g_N(1S)=g_N\left[1-\tfrac{1}{2}(Z\alpha)^2\left(\frac{m}{m+M}\right)^2\left(\frac{1+\frac{1}{3}a_N}{1+a_N}\right)+(6Z)^{-1}(Z\alpha)^2\left(\frac{m}{m+M}\right)\left(\frac{1+2a_N}{1+a_N}\right)-(3Z)^{-1}(Z\alpha)^2\left(\frac{M}{m+M}\right)\right] \tag{17}$$

where a_e and a_N are the anomalous magnetic moments of the free electron and the free nucleus and g_e and g_N are the total moments. These are in units of a Bohr magneton and a nuclear magneton respectively.

Using these two expressions we obtain the ratio of the electron to the proton g factors in the ground state of hydrogen. We have

$$\frac{g_e(H)}{g_p(H)}\cong\frac{g_e}{g_p}\left[1+\frac{\alpha^3}{4\pi}-\tfrac{1}{6}\alpha^2\frac{m}{M}\left(\frac{a_p}{1+a_p}\right)\right]. \tag{18}$$

This gives a numerical value of $(g_e/g_p)[1+2.8\times10^{-8}]$.

Equations (16) and (17) may also be used to calculate ratios of electron g factors in various hydrogenic atoms. In previous work [1, 2] the coefficient of the $(Z\alpha)^2m/M$ term in the electron g factor had been 2/3; however as discussed in this paper and in ref. [7] this is modified to 1/2. This has interesting consequences in that it changes the prediction of $g_e(H)/g_e(D)$ from $1+9.7\times10^{-9}$ to $1+7.3\times10^{-9}$. Whereas the former number agreed quite well with the experimental result $1+(9.4\pm1.4)\times10^{-9}$ given by Larson, Ramsey, and Valberg [10], the new result no longer is consistent with this determination. It is, however, compatible with the value $1+(7.2\pm3.0)\times10^{-9}$ obtained by Hughes and Robinson [11].

5. Acknowledgment

The author wishes to acknowledge useful communications with Professors S. J. Brodsky and R. A. Hegstrom.

6. References

[1] Hegstrom, R. A., Phys. Rev. **184**, 17 (1969) and Phys. Rev. **A1**, 536 (1970).

[2] Grotch, H., Phys. Rev. Letters **24,** 39 (1970).
[3] See Hegstrom, R. A., these Proceedings.
[4] Barker, W. A., and Glover, F. N., Phys. Rev. **99,** 317 (1955).
[5] Grotch, H., and Yennie, D. R., Revs. Mod. Phys. **41,** 350 (1968).
[6] Brodsky, S. J., and Primack, J. R., Ann. Phys. (N.Y.) **52,** 315 (1969).
[7] Grotch, H., Phys. Rev. **A2,** 1605 (1970). In this paper eq. (9) of the present work is given. This corrects some omissions made in Refs. [1] and [2].
[8] Foldy, L. L., and Wouthuysen, S. A., Phys. Rev. **78,** 29 (1950).
[9] Lamb, W. E., Jr., Phys. Rev. **85,** 259 (1952).
[10] Larson, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23,** 1369 (1969).
[11] Hughes, W. M., and Robinson, H. G., Phys. Rev. Letters **23,** 1209 (1969).

Ratio of Atomic g_J Values for Hydrogen and Deuterium: An Isotope Effect*

H. G. Robinson and William M. Hughes**

Physics Department, Duke University, Durham, N. C. 27706

An experimental determination of the $g_J(H)/g_J(D)$ ratio is reviewed. Optical pumping with electron spin-exchange detection of H and D Zeeman resonances was used. The applied magnetic field was ~50 G. A discussion of various possible sources of systematic error is given. Recent work indicates that the previous estimate of possible systematic error was too large. The result without estimate of systematic error is $g_J(H)/g_J(D) = 1 + (7.2 \pm 1.2) \times 10^{-9}$. This is compared with theory.

Key words: Atomic binding corrections; (experimental) electron magnetic moment; gyromagnetic ratio.

Historically speaking, there has been a continuing interest in the determination of atomic g_J values. In particular, the ratio of g_J values for hydrogen isotopes in the $^2S_{1/2}$ ground electronic state has been investigated over a period of years by workers using various experimental techniques. Beginning with the year 1949, a plot of precision achieved as a function of time is presented in figure 1. Prior to 1969, all determinations shown [1–6] were consistent with unity. The experimental determination reported on in this paper is unique in first determining the $g_J(H)/g_J(D)$ ratio which gives a value different from unity [7] and in achieving a precision in Zeeman effect measurement of ~1 part in 10^9 while using an applied magnetic field of $\lesssim 50$ G.

The method for polarizing and detecting H and D

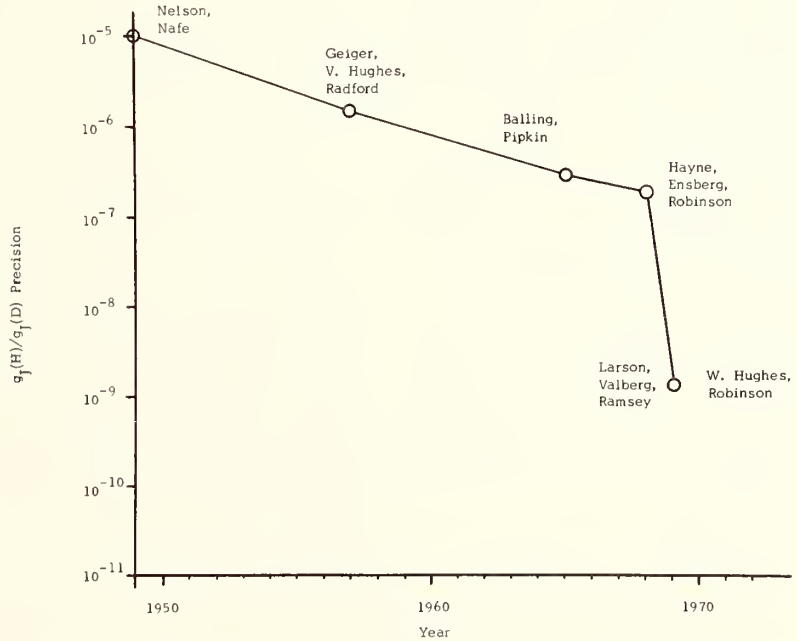


FIGURE 1. *Precision of* $g_J(H)/g_J(D)$ *determinations as a function of time.*
Prior to 1969, all values for the ratio were consistent with unity. Balling and Pipkin actually determined $g_J(H)/g_J(T)$. Straight line sections between points are arbitrarily drawn.

* Work supported in part by National Science Foundation and U. S. Office of Naval Research.
** Present address, Goddard Space Flight Center, Code 521, Greenbelt, Maryland 20770.

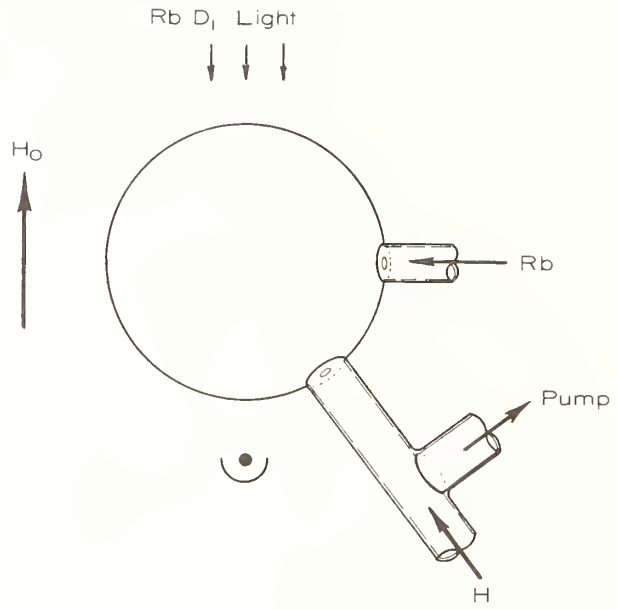


FIGURE 2. *Schematic of typical* Rb-H-D *cell.*
Various orientations of H and Rb input orifices were used. No wall coating was applied to input tubes.

Zeeman resonances used electron spin exchange with optically pumped Rb. Evacuated wall-coated spherical cells (200 cm^3) were used to obtain long relaxation times. A schematic diagram of a cell is shown in figure 2. Linewidths caused by gradients in the applied magnetic field and the first order Doppler shift were suppressed by the random rattling of the atom species in the cell. A shielded solenoid with several current shims for reducing inherent inhomogeneity was used to produce the 50 G applied magnetic field. The short-term (1 s) stability of the raw current supply driving the solenoid was ~2 parts in 10^8. The estimated H Zeeman resonance line width due to field gradients was $\lesssim$0.2 Hz. Typical H resonance line widths were ~20 Hz, with roughly equal contributions due to geometric lifetime, wall relaxation, and spin exchange. The magnetic field was locked to a frequency standard by means of a ^{87}Rb magnetometer independent of the H–D–Rb cell. Field modulation was not used. Examination of all resonances without interaction with the field stabilization loop was therefore possible. An example of the time stability achieved is shown in figure 3. Typically, the field drifted ~1 part in 10^9 during the time required to determine a pair of Zeeman resonance transitions in ^{87}Rb and H (or D). Certain effects due to long-term drift were suppressed by repeatedly measuring Rb and H alternatively. From a resonance frequency, the product of g_J and the applied field, H_0, was calculated using the Breit-Rabi equation. The ratio of $g_J(\mathrm{Rb})/g_J(\mathrm{H})$ was then formed from the appropriate $g_J H_0$ product pairs. By using different Rb pumping light intensities, the $g_J(\mathrm{Rb})/g_J(\mathrm{H})$ ratio was extrapolated to zero intensity for both senses of circular polarization. Figure 4 shows one example of such an extrapolation. Similar extrapolations give $g_J(\mathrm{Rb})/g_J(\mathrm{D})$. The summary of data is shown in figure 5. The ratio of the weighted averages for $g_J(\mathrm{Rb})/g_J(\mathrm{H})$ and $g_J(\mathrm{Rb})/g_J(\mathrm{D})$ yields [8]

$$g_J(\mathrm{H})/g_J(\mathrm{D}) = 1 + (7.2 \pm 1.2) \times 10^{-9}.$$

The error is the weighted rms deviation and represents the scatter in the results for various runs indicated in figure 5. Variation of experimental parameters included Rb pumping light intensity and

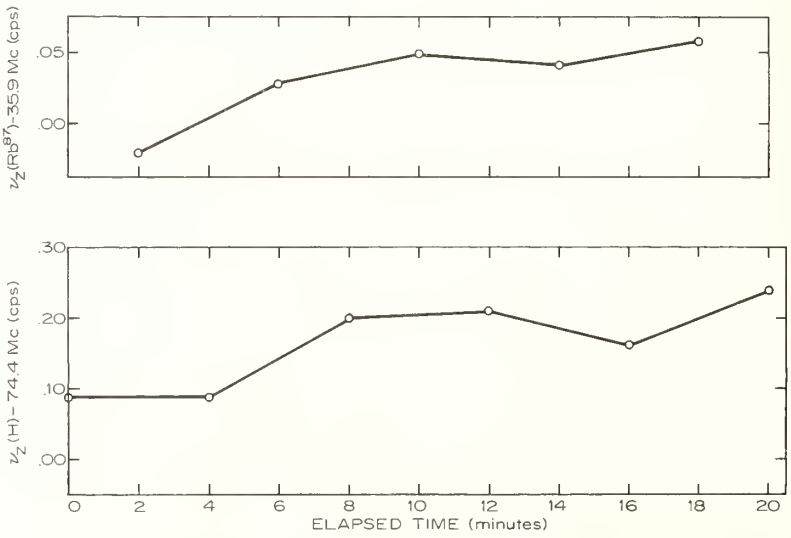


FIGURE 3. *Representative time stability of the 50 G applied magnetic field.*
The ordinates are in units of cps showing the change in Zeeman frequency ν_Z. Transition frequencies in H and ^{87}Rb were determined alternatively at times indicated by open circles. Connection by straight lines is arbitrary.

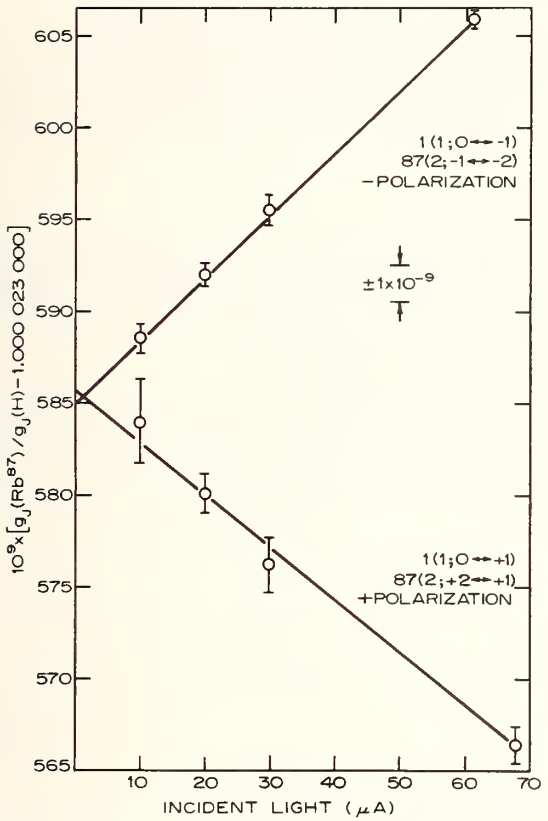


FIGURE 4. *Typical straight-line extrapolation of* $g_J(^{87}Rb)/g_J(H)$ *to zero* Rb *pumping light intensity at 50 G applied magnetic field.*

Zeeman transitions used are indicated by $(F; m_F \leftrightarrow m_F \pm 1)$.

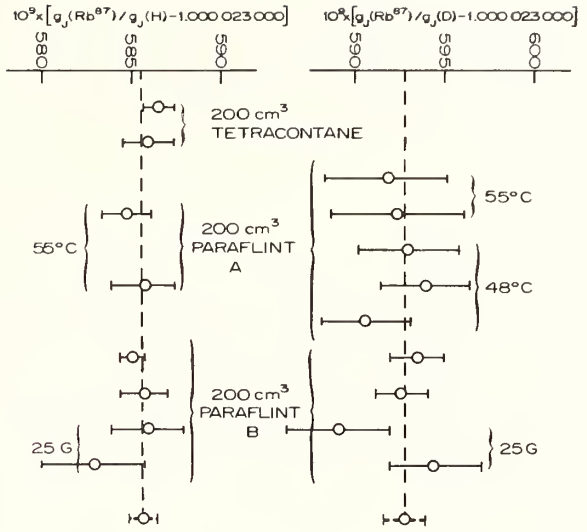


FIGURE 5. *Summary of* $g_J(^{87}Rb)/g_J(H)$ *and* $g_J(^{87}Rb)/g_J(D)$ *results.*

Data were taken at an applied field of 50 G and at a cell temperature of 35 °C unless otherwise noted. Dotted lines refer to the weighted average and weighted rms deviation of all data.

polarization, atom densities, cell temperature, magnitude and homogeneity of the applied magnetic field, type of wall coating, rf coil geometry, rf field strength, and cell geometry. Null effects have been found for Doppler shift, and spin-exchange shifts in the extrapolated results. In reference [5], we increased the error to 3.0 parts in 10^9 to take into account possible effects due to asymmetry in line shape not explicitly checked. Although a majority of the original data were taken in such a way that assessment of line-symmetry is now impossible, additional experimental work indicates the frequency shift ascribable to asymmetry in the Lorentzian line shape is $\lesssim 1$ part in 10^9; formation of the ratio of g_J-values may further suppress the effect of asymmetry. The conclusion is that the rms error of 1.2×10^{-9} is a reasonable error for the data now at hand and that we have little or no basis on which to assess remaining systematic error. We would therefore use this smaller error, emphasizing that no systematic error estimate is included. In comparing with theory [9, 10], in reference [5] it was pointed out that "experimental and theoretical values overlap only because of the increase in the error limit due to estimated systematic error." The use of the smaller error implies disagreement with this theory. However, a recent change [11] in the theoretical calculation produces agreement with our result. In comparison with the experimental result of Larson et al. [6], $1+(9.4\pm 1.4)\times 10^{-9}$, overlap in the two experimental determinations occurs. The new theoretical value lies just below their result.

The use of the usual Breit-Rabi equation for data reduction should impose no restriction on error at the low magnetic field used. An estimate of the change in the g-value of H due to configuration mixing with the $^2D_{3/2}$ state gives ~ 4 parts in 10^{16}. For a motionally narrowed Zeeman resonance in a single region spherical cell, an independent theoretical study [12] has been made of the effects caused by field gradients on line width, line shape, and frequency shift. This work supports our conclusion of no significant systematic error due to field gradients. The low magnetic field used in this experiment has been helpful in this respect, since widths and shifts are dependent on the absolute (not fractional) size of the gradient.

Interest continues in verifying the result to ~ 1 part in 10^9 and in pressing on toward higher precision for further interaction with theory. From the experience gained with this experiment, we conclude that further increases in precision are certainly feasible.

References

[1] Nelson, E. B., and Nafe, J. E., Phys. Rev. **76**, 1858 (1949).

[2] Geiger, J. S., Hughes, V. W., and Radford, H. E., Phys. Rev. **105**, 183 (1957).

[3] Balling, L. C., and Pipkin, F. M., Phys. Rev. **139**, A19 (1965). The ratio $g_J(H)/g_J(T)$ was found with a result consistent with unity.

[4] Hayne, G. S., Ensberg, E. S., and Robinson, H. G., Phys. Rev. **171**, 20 (1968).

[5] Hughes, W. M., and Robinson, H. G., Phys. Rev. Letters **23**, 1209 (1969).

[6] Larson, D. J., Valberg, P. A., and Ramsey, N. F., Phys. Rev. Letters **23**, 1369 (1969).
[7] Hughes, W. M., and Robinson, H. G., Bull. Am. Phys. Soc. **14**, 953 (1969).
[8] This result and the constants used have been given in reference [5].
[9] Hegstrom, R. A., Phys. Rev. **184**, 17 (1969).
[10] Grotch, H., Phys. Rev. Letters **24**, 39 (1970).
[11] Grotch, H., and Hegstrom, R. A., private communications. See also these Proceedings.
[12] Watanabe, S. F., Hayne, G. S., and Robinson, H. G., Bull. Am. Phys. Soc. **14**, 524 (1969) and **15**, 561 (1970).

DISCUSSION

V. W. Cohen: Were there any pressure effects that you noticed?

H. G. Robinson: Well, these are all done at essentially vacuum with the cell evacuated continuously. There were no pressure effects that we noticed. We tried to estimate the pressure that would exist in the cell at the time of running, and calculations indicate that it has virtually no effect on the averaging process or anything else. The wall is the thing one worries about.

V. W. Cohen: What was the pressure of hydrogen?

H. G. Robinson: I think probably no more than 10^{-5} torr. Bill, does that sound right? Perhaps less.

W. M. Hughes: It may have been as high as 10^{-4} torr.

V. W. Cohen: For H_2.

H. G. Robinson: That would be H_2. It's the wall that one worries about in changing, for instance, the hyperfine separation. We have thought about that and believe that it's probably all right here.

LEPTON *g*-FACTOR ANOMALIES

Methods for Lepton *g*-Factor Anomaly Measurements

G. Gräff

University of Mainz, Mainz, Germany

The experimental methods are described which so far have been used in lepton *g*-factor measurements. In particular, the advantages and disadvantages of the different methods are compared and the question of whether the accuracy of these experiments can be improved is discussed.

Key words: Anomalous magnetic moment of leptons; lepton *g*-factors.

1. Introduction

As the different tests of quantum electrodynamic theory and the meaning of anomalous magnetic moments of leptons within this theory have been discussed by Dr. Brodsky [1] and Dr. Bailey [2], I shall restrict myself to the experimental methods which so far have been used. Most of all I will compare the advantages and disadvantages of the different methods and discuss the question of whether the accuracy of these experiments can be improved. It will be the main purpose of my talk to demonstrate that the answer to this question is strongly connected with the problem of trapping polarized charged particles.

The Dirac theory of a particle with spin $\frac{1}{2}$, interacting with an external magnetic field B, predicts for this particle a magnetic moment equal to one

Bohr magneton:

$$\mu = -g \cdot \mu_0 \cdot \mathbf{s}$$

where $\mu_0 = (e\hbar/2mc)$, m is the rest mass of the particle, and $g=2$. As a consequence of the quantization of the electromagnetic field, Schwinger [3] and Luttinger [4] predicted that the electron g-factor would be different from the Dirac value of 2. We can introduce a quantity "a", called the anomaly, that measures the deviation of g from 2:

$$\mu = -2(1+a)\mu_0 \cdot \mathbf{s}.$$

The corresponding Feynman diagram of lowest order, that is of second order in e, is the electron vertex. It represents the interaction of an electron scattered by an external potential with its own and the radiation field. If the electron is scattered by a magnetic field, the energy shift corresponding to this vertex can be represented by an additional magnetic moment. This is the above-mentioned anomaly of the magnetic moment. We find to lowest order

$a = \alpha/2\pi$ α being the fine structure constant.

The second order contains mainly two effects:

(1) emission and reabsorption of two virtual photons;
(2) the vacuum polarization correction of the propagator of the first virtual photon.

The third order contributions have been calculated recently by Brodsky and his collaborators [5]. The coefficient is estimated to be 0.5. Summarizing we have

$$a_e^{\text{th}} = (\alpha/2\pi) - 0.3285(\alpha^2/\pi^2) + 0.5(\alpha^3/\pi^3).$$

for the electron/positron.

For the muon one finds

$$a_\mu^{\text{th}} = (\alpha/2\pi) + 0.76578(\alpha^2/\pi^2) + (27 \pm 3)(\alpha^3/\pi^3).$$

2. Quantities to be Measured for the Determination of "a"

Let us consider an electron in a magnetic field. We assume the magnetic field to be homogeneous and the momentum of the particle perpendicular to the magnetic field. Now there are two completely different ways of looking at the problem:

(A) The momentum of the particle revolves with the cyclotron frequency

$$\omega_c = eB/\gamma mc, \quad \text{where} \quad \gamma = [1-(v^2/c^2)]^{-1/2}.$$

The spin precesses about B with the Larmor frequency

$$\omega_L = (1+a\gamma)\omega_c.$$

The difference between these frequencies is the frequency at which the spin rotates about the momentum

$$\omega_L - \omega_c = \omega_a = a(eB/mc) = \theta/T$$

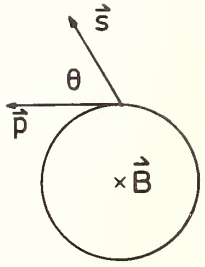


FIGURE 1a. *Dirac particle in a magnetic field* B.
The momentum revolves with frequency $\omega_c = (e/\gamma mc)B$ around B. The spin precesses around B with frequency $\omega_L = (1+a\gamma)(e/mc\gamma)B$. $\omega_L - \omega_c = \omega_a = (\theta/T)$. ($\mathbf{p} \perp \mathbf{B}$).

θ being the angle between spin and momentum, T the time. To get the anomaly "a" it is thus necessary to measure the quantities ω_a and B, assuming e/mc to be known. Since "a" is a number of order 10^{-3}, ω_a is three orders of magnitude smaller than ω_L and ω_c. Consequently a measurement of the difference frequency ω_a with the same relative accuracy as ω_L and ω_c increases the accuracy of the anomaly by three orders of magnitude. (See fig. 1a.)

(B) The energy of the particle is (in the nonrelativistic limit)

$$W/\hbar = (n+\tfrac{1}{2})\omega_c - m_s\omega_L \qquad n = 0, 1, 2, \cdots$$

$$m_s = \pm\tfrac{1}{2}$$

$$\omega_L - \omega_c = \omega_a = a(e/mc)B$$

or $a = (\omega_L - \omega_c)/\gamma\omega_c$

(See fig. 1b)

To determine the anomaly "a" the transition frequencies ω_L and ω_c have to be measured.

The selection rules are

$\Delta n = \pm 1$ for electric dipole radiation,

$\Delta m_s = \pm 1$ for magnetic dipole radiation.

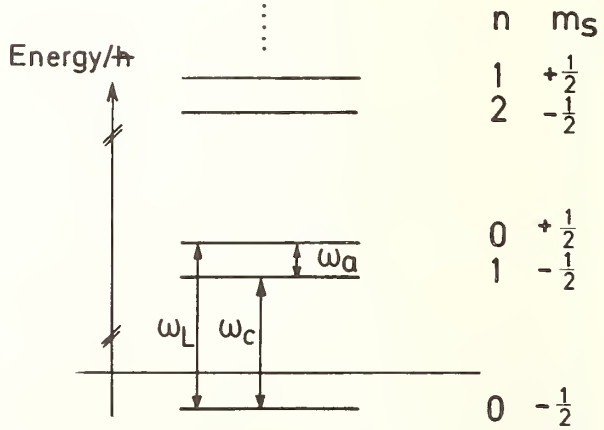


FIGURE 1b. *Schematic diagram of the energy levels of a Dirac particle in a magnetic field.*

Again the direct measurement of the transition with frequency $\omega_L - \omega_c = \omega_a$ is of main importance. This transition corresponds to a combination of a magnetic and an electric dipole transition with $\Delta n = \pm 1$, $\Delta m_s = \pm 1$ at the same time. Such a type of transition is forbidden to first order. It can be enforced by an inhomogeneous magnetic rf field.

3. General Principles

Summarizing, we see that an experiment which attempts to measure the anomalous magnetic moment of a free lepton necessarily encounters four problems:

3.1. trapping of the particle;

3.2. measurement of the trapping field either by NMR or by measuring ω_L or ω_c respectively if γ is known;

3.3. polarization of the spin of the particle;

and either

3.4.1. detection of the spin polarization vector relative to the momentum vector of the particle as a function of time in a magnetic field. Let us call this type of experiment a "geometrical" experiment;

or, alternatively,

3.4.2. the induction and detection of the relevant rf transitions ω_L and ω_c, or, if possible, ω_L or ω_c and the difference frequency ω_a directly. Let us call this type of experiment a "rf spectroscopic" experiment.

3.1. Trapping of Particles

The accuracy of the experiment is finally limited by the uncertainty relation, i.e., the time of interaction between the particle and the magnetic field. Three possibilities have been tried successfully.

3.1.1. The magnetic bottle, consisting of a homogeneous magnetic field with a superimposed relatively weak inhomogeneous magnetic field. This combination has been used in Ann Arbor by Crane and his collaborators [6, 7] to trap electrons and positrons, and also at CERN by Farley, Picasso and their group to trap muons [8]. Electrons could be trapped for times up to a few seconds.

3.1.2. The quadrupole trap, consisting of an electrostatic quadrupole with potential

$$\phi = (u/R_0^2)\,(x^2+y^2-2z^2)$$

and a superimposed homogeneous magnetic field $B = B_z$. (u = trap voltage, R_0 = trap radius.)

This combination has been used by Dehmelt [9], Farago [10], and our group at Bonn/Mainz [11]. Electrons can be trapped in these for times up to hours.

3.1.3. A third way to achieve a sufficiently long interaction time is to decelerate electrons down to the order of 10^{-5} eV. This method has successfully been used by Fairbank and collaborators [12].

3.2. Measurement of the Trapping Field

The determination of the magnetic field is of course more difficult if it is inhomogeneous. In this case one needs a map of the field and information about the mean position of the particle. These difficulties give the main contribution to the final error in the determination of the anomaly. If homogeneous magnetic fields are used, the cyclotron frequency can be induced to determine B. The gain of kinetic energy of the particle has then been observed by two different methods. One possibility is that the electrons gain so much energy that they leave the trap [11]. A different method, developed by Dehmelt [9], uses the fact that the electrons might oscillate in an LC circuit of definite resonance frequency with the same frequency. When the electrons gain energy, they transfer it partially to the circuit. Consequently the thermal noise of the circuit increases. This can be detected with high sensitivity.

It is also possible to measure the magnetic field via the Larmor frequency [13] (see below). It is important that the evaluation of B from ω_L or ω_c includes the knowledge of γ and e/mc.

3.3. Polarization of the Particles

In the geometrical experiments polarized muons are produced by the forward decay of pions [2], polarized electrons by Mott scattering [6], and polarized positrons from Beta decay [7]. In the rf spectroscopic experiments two different methods have been used. At Bonn the electrons are polarized by spin exchange with a polarized atomic beam. This was originally proposed by Dehmelt [14] and Farago [10]. Fairbank and collaborators developed a method of separating electrons in the ground state ($n = 0, m_s = -\frac{1}{2}$) in time [12]. Electrons of extremely low energy are created in pulses in a high magnetic field. Ground state electrons have a negative effective magnetic moment, whereas electrons in excited states have a positive one. As the electrons move to a lower magnetic field, ground state electrons are decelerated and electrons in excited states accelerated. By this means they can be separated in time.

3.4. Methods of Determination of the Anomaly Frequency

3.4.1. In the "geometrical" experiments the angle θ between spin vector and momentum of the particle is measured at a fixed orbital point as a function of time. The polarization of electrons is detected by Mott scattering, the polarization of positrons by exploiting the spin dependence of ortho- and para-positronium formation. The muon polarization is measured using the fact that in the rest frame the decay electrons are emitted preferentially along the spin direction (see fig. 2). As the momentum of a particle in a magnetic bottle is no longer perpendicu-

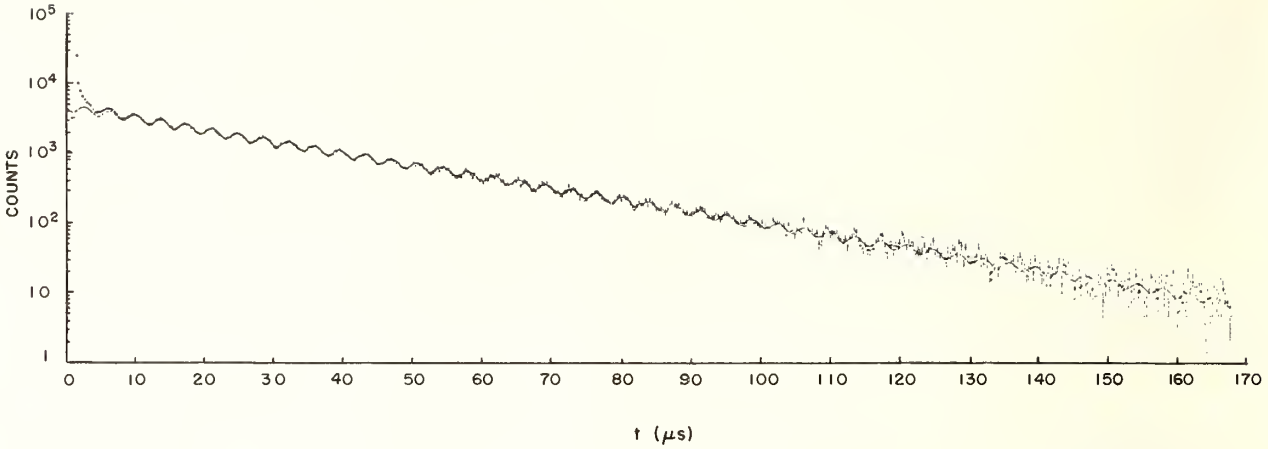


FIGURE 2. *Distribution of decay-electron events as a function of time.*
Data are fitted from 21 to 190 μs. (From ref. 2.)

lar to the magnetic field, the general formula for ω_a derived by Bargmann, Michel and Telegdi [14] has to be used. The influence of possible stray electric fields must be eliminated by measurements of ω_a at different magnetic fields.

3.4.2. In the rf spectroscopic measurements, the transition at frequency ω_a has to be induced and observed. (See fig. 3 from ref. [15].) As mentioned above, this transition is forbidden to first order and must be enforced by an inhomogeneous magnetic rf field of order gauss/cm. This field is necessarily accompanied by a homogeneous magnetic rf field. The presence of this field might lead to line shifts and line asymmetries. For the correct interpretation of ω_a the Dirac equation has to be solved for the given combination of homogeneous magnetic field and electrostatic trapping potential [15]. In the non-relativistic limit the electric field shifts the cyclotron frequency and therefore ω_a also. Hence ω_a has to be measured as a function of the electric field strength E and extrapolated to E = 0. The transition at frequency ω_a involves a jump from one cyclotron orbit to another with a spin flip at the same time. The induction of the Larmor frequency is also equivalent to a spin flip. Both transitions can therefore be detected at least in principle by the same method. One possibility is the extraction of the electrons out of the trap and a polarization analysis by Mott scattering [10]. The polarization can also be detected by the following method: Let the trapped electrons have initially an energy of about 10 eV. Then the electrons lose their energy by inelastic collisions with a beam of polarized atoms. The cross section for the inelastic electron-atom collision depends on the relative spin orientation: The cross section for triplet collisions is significantly lower than for singlet collisions. Hence an energy analysis of the trapped electrons after a certain number of inelastic collisions allows a polarization measurement and by this means the detection of ω_a and ω_L. The detection of a spin flip (ω_L) and of the cyclotron frequency ω_c is also possible by a time of flight measurement if the electron was initially in the ground state [12].

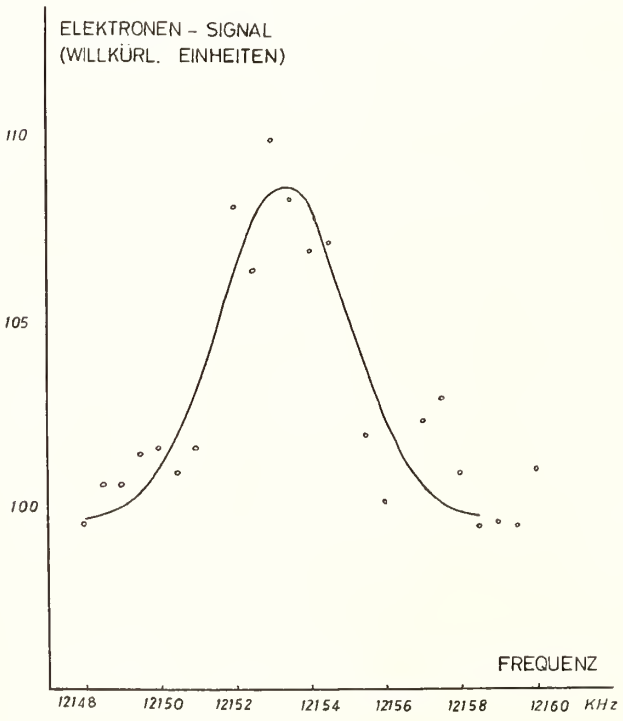


FIGURE 3. RF-*spectroscopically induced anomaly frequency* ω_a.
(From ref. [15].)

4. Summary

All the g-2 experiments performed measure the interaction between the magnetic moment of the particle and a homogeneous magnetic field superimposed by an inhomogeneous magnetic or electric trapping field. The trapping fields limit the accuracy of the experiments. The accuracy can be increased

by decreasing the relative inhomogeneity. A corresponding experiment has been performed at Ann Arbor [16] and is under way at Mainz. For technical reasons a further substantial increase of the magnetic field with reduction of the relative inhomogeneity is not possible in the g-2 experiment of the muons. Therefore it has been proposed by Bailey, Farley and Picasso at CERN to replace the weak focussing magnetic system by electrostatic quadrupoles [17]. Then the cyclotron frequency will be shifted by the applied electric field, and the Larmor frequency by the motional magnetic field. As the latter depends on $\gamma=[1-(v^2/c^2)]^{-1/2}$, there exists a specific value γ_0 where the shift of the cyclotron frequency is equal to the shift of the spin precession frequency. Hence at this energy the rate of change of the angle between spin and momentum does not depend on the electric field strength. For muons this "magic" value is $\gamma_0=29.3$. This corresponds to an energy of about 3 GeV for muons and 14 MeV for electrons. It retains the disadvantage that ω_c and ω_a have to be determined by different methods.

The rf spectroscopic methods can work with low energy particles. They have the advantage that ω_a and ω_L can be measured under identical conditions using the same method. The main disadvantages are that the transition at frequency ω_a is forbidden to first order, and the presence of unwanted homogeneous magnetic rf fields. At present the relative line width is of the order 10^{-4}, but line width and line shape are still not fully understood. The accuracy of the electron experiment performed by Rich and Wesley at Ann Arbor is therefore at present an order of magnitude higher than the accuracy of any other experiment.

5. References

[1] Brodsky, S. J., these Proceedings.
[2] Bailey, J., these Proceedings. See also E. Picasso: Current Developments in the Study of Electromagnetic Properties of Muons, International Conference on High-Energy Physics and Nuclear Structure, Columbia University, New York, 1969.
[3] Schwinger, J., Phys. Rev. **73**, 416 (1948).
[4] Luttinger, J., Phys. Rev. **74**, 893 (1948).
[5] Aldins, J., Brodsky, S. J., Dufner, A. J., and Kinoshita, T., Phys. Rev. Letters **23**, 441 (1969).
[6] Wilkinson, D. T., and Crane, H. R., Phys. Rev. **130**, 852 (1963).
[7] Gilleland, J., and Rich, A., Phys. Rev. Letters **23**, 1130 (1969).
[8] Bailey, J., et al., Phys. Letters **28** B, 287 (1968).
[9] Dehmelt, H. G., and Walls, F. L., Phys. Rev. Letters **21**, 127 (1968).
[10] Farago, P. S., Advances in Electronics and Electron Physics **21**, 1 (1965).
[11] Gräff, G., und Klempt, E., Z. Naturforschung **22a**, 1960 (1967).
[12] Kincaid, B., Fairbank, W. M., and Knight, L. V., this Conference.
[13] Gräff, G., Major, F. G., Roeder, R. W. H., and Werth, G., Phys. Rev. Letters **21**, 340 (1968).
[14] Bargmann, V., Michel, L., and Telegdi, V., Phys. Rev. Letters **2**, 435 (1959).
[15] Gräff, G., Klempt, E., and Werth, G., Z. Physik **222**, 201 (1969).
[16] Wesley, J. C., and Rich, A., Phys. Rev. Letters **24**, 1320 (1970).
[17] Bailey, J., et al., CERN—Report PHI/COM—1969.

DISCUSSION

B. M. Kincaid: What are you doing about the line shape? Do you have a theory about it yet?

G. Graff: We have not a theory for the line shape, at least not an exact theory for the line shape, but we know that the line shape does depend upon the rf amplitude, the amplitude of the inhomogeneous rf, so what we do at present is to measure the dependence of the line shape upon the rf amplitude.

B. M. Kincaid: Then extrapolate to zero amplitude?

G. Graff: Yes. That is what we probably will do. But mainly we try to go to higher fields to achieve from the very beginning a smaller line.

B. M. Kincaid: Would you hazard a guess as to your final accuracy?

G. Graff: I think that the final accuracy is probably limited by the energy of the trapped particles. The energy is of the order of, say, a few eV. In units of the rest mass, it's a few times 10^{-6}. So I assume that the final limitation of the accuracy of the rf spectroscopic experiments will be of the order of parts per million.

B. M. Kincaid: On g minus 2?

G. Graff: On g minus 2, yes, since this is a real g minus 2 experiment, not a g experiment.

Preliminary Results of a New Measurement of the *g*-Factor Anomaly of the Free Electron*

John C. Wesley and Arthur Rich

The Harrison M. Randall Laboratory of Physics, The University of Michigan, Ann Arbor, Mich. 48104

We report a new measurement of the electron *g*-factor anomaly. Our preliminary result is $a=(1\ 159\ 644\pm7)\times10^{-9}$, in good agreement with current QED theoretical predictions. We discuss the role of the constant μ_p'/μ_B (the magnetic moment of the proton in Bohr magnetons, measured in a water sample) in our experiment. The NMR techniques used to measure the magnetic field are detailed.

Key words: Electron *g*-factor anomaly; proton magnetic moment in Bohr magnetons; quantum electrodynamics (comparison to theory).

1. Introduction

We are in the process of completing a new determination of the *g*-factor anomaly of the free electron. A preliminary discussion of the results of this effort has been published elsewhere [1]. In this paper we will emphasize those aspects of our work most closely related to the fundamental constants. We begin with a brief description of the experiment itself. This is followed by a discussion of the significance of the electron *g*-factor as a fundamental constant. The third section of this paper contains a detailed consideration of the role of the constant μ_p'/μ_B in our current measurement. The concluding section summarizes our preliminary results and discusses prospects for future improvements in accuracy.

2. Experimental Technique

The Michigan technique for measuring *g*-2 involves direct observation of the precession of the electron spin in a magnetic field. Partially polarized electrons obtained by Mott scattering are injected into a weak magnetic mirror trap (mirror ratio 1.0001). After being confined for an accurately measured time T, they are ejected from the trap to undergo a second Mott scattering. Those which scatter through approximately 90° are counted.

While the electrons are trapped, their average spin motion may be described as a precession of their polarization $\mathbf{P}$, relative to their velocity $\mathbf{v}$, at a frequency ω_D. If $\mathbf{v}$ is perpendicular to a uniform field $\mathbf{B}$ and no electric fields are present, ω_D is given by

$$\omega_D = a\gamma\omega_c = a\omega_0, \tag{1}$$

where $\omega_0 = eB/m_0c$ and $\gamma = (1-v^2/c^2)^{-1/2}$.

The number of electrons counted after the second scattering varies as $\mathbf{P}\cdot\mathbf{v}$. If the number of electrons ejected is independent of T, the counting rate is given by

$$R(T) = R_0[1+\delta\cos(\omega_D T+\phi)], \tag{2}$$

where $R(T)$ is the rate for a polarized electron beam, R_0 is the rate for an unpolarized beam, δ is the Mott asymmetry factor (typically 0.05), and ϕ is a phase constant. If n oscillations of $R(T)$ are observed between $T=T_1$ and $T=T_2$, then $\omega_D = 2\pi n/(T_2-T_1)$.

The zero energy cyclotron frequency, ω_0, is measured indirectly by using the relation $\omega_0 = (\mu_p'/\mu_B)^{-1}\omega_p(H_2O)$, where μ_p'/μ_B is the proton magnetic moment in a water sample, in Bohr magnetons, and $\omega_p(H_2O)$ is the resonance frequency of protons in a water sample. (We discuss this conversion at greater length in a following section.)

Figure 1 shows a diagram of the experimental region. The entire area is immersed in a nearly homogeneous magnetic field produced by a wire-wound solenoid of 2 ft diam and 8 ft length (not shown). The variation of the magnetic field as a function of the axial coordinate z that forms the mirror trap is shown below the diagram. The nominal field of 1.2 kG and electron energy of 110 keV result in a 1.0 cm cyclotron radius. Voltage pulses on the inject and eject cylinder serve to inject and eject the electrons at the desired times.

The asymmetry curve of eq (2) is obtained by sampling $R(T)$ at various values of T and fitting the data by least squares methods to yield the position of a maximum at both an inner (1500 μs) and an outer (4000 μs) trapping time. Several additional curves taken between these two maxima permit the determination of the number of cycles separating them, typically about 10^4. Since the position of the maxima can be established to a small fraction of a cycle, ω_D can be measured to a few ppm.

* This work supported by the U. S. Atomic Energy Commission.

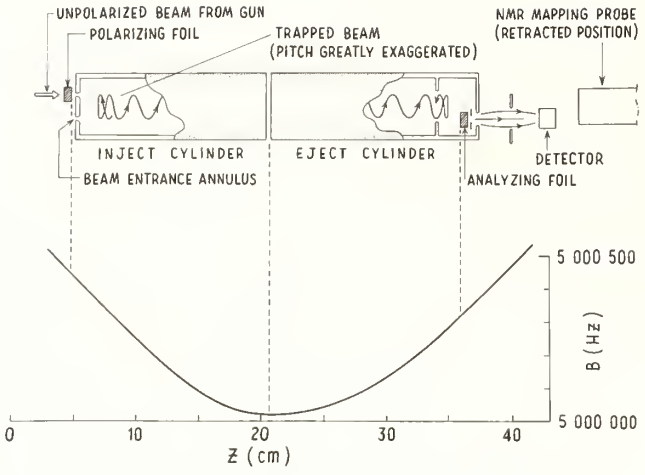


FIGURE 1. *Components in the g-IV trapping region.* The variation of the magnetic field as a function of the axial coordinate z is shown below the diagram.

The magnetic field is controlled by an NMR servo-regulator. The field in the region where the electrons are trapped is measured with a second NMR probe. The entire trapping assembly may be rotated aside and the mapping probe extended while the system is evacuated, thus permitting field measurements immediately before and after an asymmetry data run.

We note that eq (1) requires corrections of the order of 10 ppm to account for the inhomogeneity of the field over the trapping region, the finite axial velocity of the trapped beam, the distribution of various axial energies of the trapped electrons, and the presence of stray electric fields in the trapping region. These considerations, as well as complete details of present and past g-factor measurements, are discussed in references [1–3].

The major uncertainties in the Michigan g-2 technique are due to (1) statistical error in ω_D, (2) uncertainty in the magnetic field experienced by the electrons due to the inhomogeneities necessary to trap the beam, and (3) systematic shifts in ω_D due to the presence of stray electric fields in the trapping region. The increased accuracy of our present experiment over the work of Wilkinson and Crane [3] is due to an increase of the magnetic field from 150 to 1200 G, in combination with a simultaneous reduction in the relative depth of the magnetic well (mirror ratio) from 1.001 to 1.0001.

3. The Electron g-Factor as a Fundamental Constant

If we substitute $S=\hbar/2$ and $g=2(1+a)$ into the equation relating magnetic moment to spin, $\mu_e=(g/2)(e/m_0c)S$, we obtain

$$\mu_e/\mu_B=(1+a), \tag{3}$$

where $\mu_B=e\hbar/2m_0c$ is the Bohr magneton. Since a is approximately 10^{-3}, the uncertainty in μ_e/μ_B is $1:10^3$ that of a itself. Our measurement of a to 6 ppm therefore defines μ_e/μ_B to 0.007 ppm. The magnetic moment of the electron in Bohr magnetons is thus the most accurately measured property of an elementary particle. It is known approximately one thousand times more accurately than either e, m_0, or e/m_0.

The most direct metrological use of μ_e/μ_B is as an auxiliary constant (one whose uncertainty is sufficiently small so that it may be considered exactly known) in least squares adjustments of the fundamental constants. A typical example of the use of μ_e/μ_B occurs in interpreting the data of the experiment of Lambe and Dicke [4] to obtain μ_p'/μ_B. Lambe and Dicke measured $g_j(\mathrm{H})/g_p(\mathrm{H_2O})$, where $g_j(\mathrm{H})$ is the electron g-factor in atomic hydrogen. The free electron g-factor g is obtained using $g_j(\mathrm{H})=g(1-\frac{1}{3}\alpha^2)$ to yield the ratio $g/g_p(\mathrm{H_2O})$. Since $g/g_p(\mathrm{H_2O})=\mu_e/\mu_p'$, we may write

$$\mu_p'/\mu_B=[g/g_p(\mathrm{H_2O})^{-1}](\mu_e/\mu_B).$$

Although the Lambe-Dicke measurement is extremely accurate ($\sigma=0.06$ ppm), the additional conversion using μ_e/μ_B introduced no significant additional error. A similar conversion applied to the data from the hydrogen maser measurements of Myint et al. [5] yields a value for μ_p/μ_B (μ_p is the magnetic moment of the free proton).

Although the electron g-factor is useful in metrology, as noted above, its principal importance currently lies in testing the validity of quantum electrodynamics (QED). Perturbation theory techniques can be used to calculate a as a power series expansion in the fine structure constant α. The most recent theoretical estimate of the sixth-order (α^3) coefficient is 0.39 [6, 7], but this is not an exact result, since many Feynman diagrams are only estimated. To our knowledge, attempts to calculate the coefficient exactly are currently underway at the University of Texas [8] and the University of Michigan [9]. Using the value of 0.39 for the sixth-order coefficient, the anomaly is given by:

$$a_{\mathrm{theo}}(e^-)=0.5(\alpha/\pi)-0.32848(\alpha/\pi)^2+0.39(\alpha/\pi)^3. \tag{4}$$

Experimental measurements of a can be interpreted as either (1) testing QED, assuming a value of α derived without QED theory, or (2), providing a QED measurement of α, assuming the validity of the theoretical calculation for a. Since recent measurements using the ac Josephson effect yield accuracies of better than 2 ppm in α, with an additional increase in accuracy possible in the near future, the first interpretation appears to be the most fruitful. Hopefully, the results of the exact sixth-order calculations for a will soon be available. The term $(\alpha/\pi)^3$ is 11 ppm of a. If the coefficient of the sixth-order term is of order unity, an experimental precision of several ppm will be necessary to verify it to a reasonable accuracy.

4. The role of μ_p'/μ_B in the Electron g-Factor Experiment

Several recent measurements of μ_p'/μ_B are of critical importance in our g-2 measurement. To obtain the zero-energy cyclotron frequency from the proton resonance frequency, we use the relation

$$\omega_0 = (\mu_p'/\mu_B)^{-1}\omega_p(H_2O),$$

since by definition,

$$\frac{\mu_p'}{\mu_B} = \frac{2\mu_p' B/\hbar}{2\mu_B B/\hbar} = \frac{\omega_p(H_2O)}{\omega_0}.$$

In the experiment of Wilkinson and Crane, the results of Liebes and Franken [10] were used for this conversion. The 7 ppm uncertainty introduced from the conversion was small in comparison with the 30 ppm error from other factors. For more precise work, this error would soon become the limiting factor. Fortunately, work by Lambe and Dicke [4], and Klein [11], each using separate techniques, has yielded measurements of μ_p'/μ_B to accuracies of 0.5 ppm or better. These results agree with each other to 0.2 ppm, well within their assigned error.

As we have discussed in the preceding section, the Lambe-Dicke experiment measured $g_j(H)/g_p(H_2O)$ to ± 0.06 ppm. In order to obtain μ_p'/μ_B, one must first correct to the free electron g-factor and then convert using μ_e/μ_B. This conversion introduces no appreciable error in our work, since μ_e/μ_B is relatively independent of the precise value of a. Specifically, if a is taken as known to ± 100 ppm from the Wilkinson-Crane experiment, the error in μ_e/μ_B is ± 0.1 ppm, and μ_p'/μ_B can be found to the same accuracy. To employ the Lambe-Dicke result in our NMR field measuring technique (using a marginal oscillator), we would have to introduce an additional correction for the effect of the paramagnetic ion doping added to the water sample. This would increase the total uncertainty in the $\omega_p(H_2O)$ to ω_0 conversion to ± 0.2 ppm.

Klein measured the ratio of the resonant frequencies of free electrons to that of protons in water. The water sample was a 0.2 M solution of $CuSO_4$ in a long ($L/D > 10$) cylinder oriented perpendicular to the magnetic field. The proton frequency was determined by exciting the probe with a marginal oscillator. Klein's result was [12]

$$[\omega_p(s)/\omega_0]_K = 1.5209945 \times 10^{-3} \ (\pm 0.5 \text{ ppm})$$

where $\omega_p(s)$ is the resonance frequency of the standard probe. We have chosen to use this result instead of the Lambe-Dicke value (in spite of the somewhat larger uncertainty of Klein's number). This choice was made because of the ease with which Klein's standard probe may be reproduced and used, as well as the fact that his experimental measurement is directly related to the conversion we wish to make.

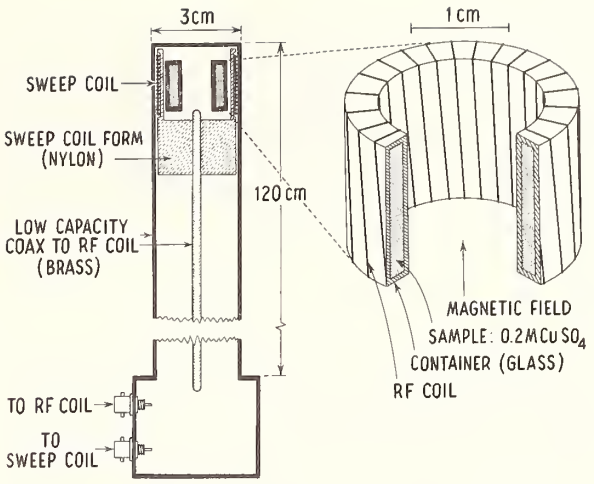


FIGURE 2. *The NMR mapping probe.*

In order to use Klein's result, we need only measure the magnetic field using the standard probe. To do this accurately, it is necessary that the probe (1) respond only to the field in the region of (1.0 ± 0.1) cm radius in which the electrons are trapped, and (2) be capable of fitting inside the vacuum tank of the assembled experiment. The standard probe is too large physically to satisfy either requirement.

The probe used to map the field has a toroidal configuration, as shown in figure (2). The water sample has the same radius and thickness as the trapped electron beam, so it samples the same magnetic field as that experienced by the electrons. The variation of the magnetic field over the probe volume is small (<5 ppm) compared to the natural line width (25 ppm) of the 0.2 M $CuSO_4$ solution. Consequently, the resonant frequency of the probe represents, to an excellent approximation, the spatial average of the field over the sample volume. The probe automatically gives the average of the magnetic field over the electron orbit. A lock-in amplifier and varicap diode was used in a servo loop to lock the marginal oscillator frequency to the magnetic field. The resonant frequency is read directly from a frequency counter. Before this frequency may be converted to ω_0, it is necessary to apply a correction for the bulk diamagnetic properties of the water-paramagnetic salt solution, as well as for possible magnetic contamination of the materials used to construct the probe.

The bulk diamagnetic correction could in principle be calculated for the toroidal configuration of our mapper probe. However, it is simpler to measure this correction, as well as that from possible magnetic contamination of the probe materials, by comparing the mapper probe with the standard probe in the same magnetic field. The field at the center of the main solenoid is homogeneous to better than 0.1 ppm over the probe volumes, and is stabilized to 0.5 ppm over times of the order of one minute. By repeatedly interchanging the mapper and standard probes in the

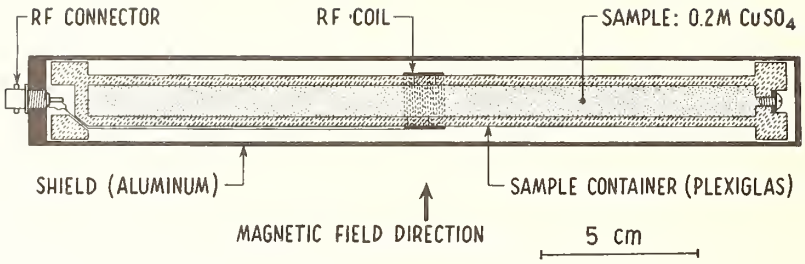


FIGURE 3. *The standard NMR probe.*

same field region, we are able to determine the ratio $[\omega_p(M)/\omega_p(s)]_{cal}$ to 0.4 ppm, where $\omega_p(M)$ is the frequency of the mapping probe. This procedure allows converting the mapper frequency to that of a standard probe at the same location.

The effect of magnetic contamination in the materials used to construct the standard probe poses a potentially serious problem. If an object of permeability μ is immersed in a uniform field, the field near the object will be altered by a factor of order $\mu-1$. Since many "nonmagnetic" substances such as aluminum, copper, glass, and plastic have $\mu-1$ in the range 10^{-5} to 10^{-6}, this shift in field is large for measurements to better than 1 ppm. We have tested the various materials used near the trapped beam or in the probes using a pair of NMR probes. The object to be tested is brought near one of the probes, and any relative shift in the field between the probes is noted. The technique can detect effects as small as 0.1 ppm. We have found that almost all materials have a detectable effect, usually of order 1 ppm. Some samples of brass shift the field by up to 10 ppm.

The effect of the materials used to construct the probe may be minimized by the proper choice of configuration. If a long hollow cylinder of permeability μ is placed perpendicular to a uniform field, the field inside is altered by a factor of order $(\mu-1)^2$. If the standard probe is constructed in the form of long coaxial cylinders and placed perpendicular to the field, the materials will have no detectable effect for $(\mu-1)<10^{-4}$, so long as the materials have uniform permeability. The effect of possible nonuniform magnetic contamination in the materials must still be considered. This was tested for by fabricating three standard probes from the same materials and checking one against the other. No significant shift was detected. Figure 3 shows a diagram of the standard probe.

The final conversion from the directly measured NMR frequency $\omega_p(M)$ to ω_0 is given by:

$$\omega_0=\omega_p(M)[\omega_p(s)/\omega_p(M)]_{cal}[\omega_p(s)/\omega_0]_K^{-1}. \quad (5)$$

The calibration factor for the mapper probe used in the preliminary data runs was $1-(1.3\pm0.4)\times10^{-6}$. The total error in the conversion is therefore ±0.7 ppm.

Sands [13] has brought to our attention the possibility of a nonlinear dependence of $\omega_p(H_2O)$ on the magnetic field intensity, due to the presence of excited states of the water molecule. Since most measurements of quantities leading to μ_p'/μ_B have been made at a field of approximately 3 kG, such a nonlinearity would cause an error in applying these results to our 1 kG field. Several experiments show that such an effect is less than ±2 ppm for an extrapolation from 1 to 3 kG. One estimate of the nonlinearity can be obtained by comparing the work of Geiger, Hughes and Radford [14], and Koenig, Prodell and Kusch [15], who measured $g_j(D)/g_p(\text{oil})$ at 3 kG and $g_j(H)/g_p(\text{oil})$ at 1.5 kG, respectively. Their results differ by $(0.3\pm1.7$ ppm), implying a nonlinearity of less than ±2 ppm in the extrapolation from 3 kG to 1 kG.

Huggins and Sanders have searched directly for a dependence of $\omega_p(H_2O)$ on magnetic field [16]. Their experiment compared the ratio of resonance frequencies of protons and various nuclei at fields of 23 and 10800 G. Their most accurate result was:

$$\frac{[\omega_p(H_2O)/\omega_{^{19}F}(H_2O)]_{23\text{ G}}}{[\omega_p(H_2O)/\omega_{^{19}F}(H_2O)]_{10,800\text{ G}}}=1-(1.6\pm2.4)\text{ ppm}.$$

If we assume that the nonlinearity is proportional to the field, this implies any effect in extrapolating from 3 to 1 kG is less than about ±0.8 ppm. If the nonlinearities depend quadratically on the field (as described below) the effect is less than ±0.3 ppm. (This result assumes, of course, that any nonlinear effects would be substantially different for the two nuclear species. If the nonlinearities were similar or equal, the relative effect would be small or zero.)

Ramsey has made theoretical estimates of field-dependent nuclear magnetic shielding [17]. For a nucleus in an external field H_0, the field at the nucleus is $(1-\sigma)H_0$, where the shielding coefficient σ arises chiefly from diamagnetic circulation of the orbital electrons. For water, $\sigma(H_2O)=(26.0\pm0.3)$ ppm. Ramsey has considered the possible dependence of σ on H_0 in the form $\sigma=\sigma_0+\sigma_1H_0^2$. He shows that for substances with $\sigma<10^{-4}$, the term $\sigma_1H_0^2$ is of order 10^{-9} at a field of 60 kG. For the lower fields of our work, the effect of this dependence is completely negligible. Ramsey also notes the possibility of a field-dependence of the bulk magnetic susceptibility of the solution. Since the total correction for this effect is of order 1 ppm, any small nonlinearity from this source should again have negligible effect.

A third possible nonlinear effect may arise from

the magnetic field affecting the population distribution of the molecular states. If σ_k is the shielding coefficient of the kth molecular state, the total shielding is given by

$$\sigma = \frac{\sum_k \sigma_k \exp -[(w_k - \frac{1}{2}\xi_k H_0^2)/kT]}{\sum_k \exp -[(w_k - \frac{1}{2}\xi_k H_0^2)/kT]}$$

with ξ_k is the susceptibility of the kth state, and w_k is the energy of the state [17]. Investigation of this point is continuing. However, we have tentatively concluded that the various nonlinearities discussed above cause no significant error in the magnetic field conversion in our work.

4. Current Values of the Electron and Positron g-Factors

Evaluating the power series expression for $a^{\text{theo}}(e^-)$ using $\alpha^{-1} = 137.036\ 08 \pm 0.000\ 26$ gives $a^{\text{theo}}(e^-) = (1\ 159\ 642 \pm 2) \times 10^{-9}$, where the error is due only to error in α^{-1}. The preliminary results of our current work give $a^{\text{expt}}(e^-) = (1\ 159\ 644 \pm 7) \times 10^{-9}$, in agreement with theory. We note again that the theoretical calculation contains only an estimate of the sixth-order coefficient. We hope to improve the precision of our experimental result by a factor of 2. Assuming the α^3 coefficient is of order unity, this will permit a reasonably accurate check of QED to sixth order.

A note concerning the positron g-factor may be of interest here. We have been able to measure the positron anomaly to about ± 1000 ppm [2]. The positron and electron g-factors are found to be equal to ± 1 ppm, as predicted by TCP invariance. Any possible attempt to measure $a(e^+)$ to better than 100 ppm using the Michigan g-2 method will encounter a serious metrological difficulty. Since no measurement analogous to μ_p'/μ_B exists for positrons, one is forced to use the relation

$$a(e^+) = \omega_D/(\omega_0)_{e^+} = (\omega_D/B)[m_0 c/e]_{e^+} \qquad (6)$$

to evaluate a. At present the constant $(e/m_0 c)_{e^+}$ is known to ± 71 ppm, and would therefore become a limiting factor for a more precise measurement. Of course, using TCP invariance, one could assume that $(e/m_0 c)_{e^+} = (e/m_0 c)_{e^-}$, but this would defeat the goal of using $a(e^+)$ to test TCP.

5. References

[1] Wesley, J. C., and Rich, A., Phys. Rev. Letters **24**, 1320 (1970).
[2] Gilleland, J. R., and Rich, A., Phys. Rev. Letters **23**, 1130 (1969).
[3] Wilkinson, D. T., and Crane, H. R., Phys. Rev. **130**, 852 (1963).
[4] Lambe, E. B. D., Ph.D. thesis (1959), Princeton University (unpublished).
[5] Myint, T., Kleppner, D., Ramsey, N. F., and Robinson, H. G., Phys. Rev. Letters **17**, 405 (1966).
[6] Aldins, J., Brodsky, S. J., Dufner, A. J., and Kinoshita, T., Phys. Rev. (to be published).
[7] Brodsky, S., and Kinoshita, T., Theoretical Results for Sixth Order Contributions to the Anomalous Magnetic Moment of the Muon and Electron (submitted to the XVth International Conference on High Energy Physics in Kiev).
[8] Campbell, J. W., private communication.
[9] Yao, E., private communication.
[10] Liebes, S., Jr., and Franken, P., Phys. Rev. **116**, 633 (1959).
[11] Klein, E., Z. Physik **208**, 28 (1968).
[12] Ibid., p. 55.
[13] Sands, R. H., private communication.
[14] Geiger, J. S., Hughes, V. W., and Radford, H. E., Phys. Rev. **105**, 183 (1957).
[15] Koenig, S. H., Prodell, A. G., and Kusch, P., Phys. Rev. **88**, 191 (1952).
[16] Huggins, R. W., and Sanders, J. H., Proc. Phys. Soc. **86**, 53 (1957).
[17] Ramsey, N. F., Phys. Rev. A **1**, 1320 (1970).

DISCUSSION

E. R. Cohen: What is the uncertainty with which you can determine the exact number of cycles between your two time measurements, and what is the effect of that on perhaps reconciling the g-III and g-IV? I made an estimate at one time that if you could get away with saying that there was a one-cycle error in g-III this could be about the right order of magnitude.

J. C. Wesley: Yes, this is an excellent observation. I had a slide which showed this, and I left it out in the interests of time. This is a very good point, because the one-cycle miscount contributes an error of about 120 parts per million, and this is just the right error to make the results more or less in agreement. The probability (statistical probability) of a one-cycle miscount in any one of the four points is about one part per million, about six standard deviations in our cycle counting procedure. In addition, if you plot those points and then you plot the points as they would have been with either a plus or minus one-cycle miscount, you get a series of ghost points either above or below. The number of cycles in each case was not the same. It varied by about a factor of two. So you find that the miscounted values do not form anything like a straight line, and the only semblance of a straight line through these various points is the one shown in the graph, so we are quite confident that we are not making this error.

G. R. Henry: Would it be very difficult to introduce purposely some electric fields to see if you get the proper relationship between observed frequency and the various field components?

J. C. Wesley: I would say we have verified the relationship to about 10 percent. An electric field does shift the difference frequency as we expect. It's a little bit difficult to get in there and induce fields of the proper magnitude, in that you want to avoid putting an electrode down the center of the experiment. But we have been able, for instance, to apply voltage to the trapping cylinders, and this does shift the frequency by about 20 or 30 parts per million. As far as we can calculate the electric field we are putting

on, the results seem to be in pretty good agreement with what we would expect. I should point out that no one has ever proved that it's really an electric field in there. It looks like an electric field. It tastes like an electric field. But there is no way to get in there and make an actual measurement of a field of 1 millivolt per centimeter. At least we don't know how to do it. If any of you know how, I would certainly be glad to find out how it's possible to do this. I think this is really a serious problem for many atomic physics experiments because I'm sure we are not the only people who are plagued by these electric fields, and perhaps this would be a good project to develop a reliable way of measuring fields of 1 millivolt per centimeter inside a small electrode region. I'll give that assignment to you out there in the audience.

A. H. Wapstra: If you could just double your precision, I think that you could reformulate your result as being one of the most accurate and one of the most direct determinations of the fine structure constant. What do you think about this interpretation?

J. C. Wesley: Well, I think you'd have to talk to Stan Brodsky about that. Initially, when we set out to do this experiment, there was no discrepancy between theory and experiment. The recalculation came from studies intended for this experiment. We were going to measure the fine structure constant at that time. Alpha was rather discrepant at that time. I believe you can interpret the results either way, depending on whether you are more confident in the theoretical calculations or the other measurements of the fine structure constant. It's certainly a good check of the agreement between QED and fundamental constants.

S. J. Brodsky: I agree. But, of course, as I point out in my talk, the measurement of the anomalous moment of the electron is the best test of the sixth order corrections. It's the only chance we have left if we are hoping for some sign of a discrepancy to show up in the precision tests of QED.

E. R. Cohen: In other words, you're not willing to believe that sixth order correction yet?

S. J. Brodsky: There is of course great interest in learning what the precise QED value is for the sixth order electron moment. Hopefully the final calculations will be completed within a year or two.

J. C. Wesley: We were rather disappointed when the values starting coming out around 640 or so. It just didn't work out. We thought we might become a little famous or something. (*Laughter.*)

Muon g-2 Measurement

J. Bailey

CERN, Geneva, Switzerland

The results of a recent experiment at CERN to measure the anomalous moment of the muon are summarized and compared with theory. New techniques now being implemented for improving the measurement are also noted.

Key words: Anomalous moment of the muon; muon; muon storage ring.

1. First Muon Storage Ring

A recent measurement [1] at CERN of the extra-Dirac part of the muon magnetic moment gave

$$a_\mu(e/m_\mu) = (370805 \pm 99) \times 10^{-8} \text{ (MHz/H}_2\text{O proton MHz)} \quad (1)$$

for muons in vacuum relative to protons in water.

Other people [2, 3] have measured the whole muon magnetic moment. One result [2] is

$$(1+a_\mu)(e/m_\mu) = (3183347 \pm 9) \times 10^{-6} \text{ (MHz/bare proton MHz)} \quad (2)$$

for muons relative to protons in the same environment.

Since the diamagnetic shielding of protons in water is 26 ppm ("ppm" means "parts per million"), we conclude that the so-called "anomaly" (more correctly, the extra-Dirac part of the half-g-factor) is

$$a_\mu^{\text{exp}} = (116616 \pm 31) \times 10^{-8}. \quad (3)$$

2. Theoretical Value

This is to be compared with theoretical calculations, which in July 1970 had reached the following stage. First, the QED contributions to the difference $(a_\mu - a_e)$ have been calculated completely up to sixth order, by two different groups [4, 5], giving

$$(a_\mu - a_e)^{\text{th}} = 1.09426(\alpha^2/\pi^2) + (20.3 \pm 1.3)(\alpha^3/\pi^3). \quad (4)$$

Secondly, all QED diagrams contributing to a_e, except those containing no lepton loops, have been calculated [5] up to sixth order, giving

$$a_e^{\text{th}} = (\alpha/2\pi) - 0.32848(\alpha^2/\pi^2) + (0.26 + C_3^{\text{no-loop}})(\alpha^3/\pi^3). \quad (5)$$

Also, the coefficient $C_3^{\text{no-loop}}$ has been estimated [6] to be

$$C_3^{\text{no-loop}} = 0.13. \quad (6)$$

Finally, the contribution of strong-interacting systems to a_μ has been estimated from the results of colliding e^+–e^- beams to be

$$a_\mu^{\text{hadronic}} \geq 6 \times 10^{-8}. \quad (7)$$

An interesting feature of this last number is that it is positive definite, and any future discoveries of resonances or large cross sections in the e^+–e^- system can only increase its value.

The sum of all these terms gives

$$a_\mu^{\text{th}} = 116586 \times 10^{-8}. \quad (8)$$

$$(a_\mu^{\text{exp}} - a_\mu^{\text{th}}) = (30 \pm 34) \times 10^{-8}. \quad (9)$$

3. New Muon Storage Ring

We are at present designing a new muon storage ring, and expect to measure a_μ with an error of less than 2×10^{-8}, thus testing the above theory much more stringently. This improved accuracy comes from the three following factors.

Firstly, the storage ring will be filled by injecting into it a momentum-selected pion beam matched to the ring acceptance; this will give very much less background (and, incidentally, higher muon polarization and beam intensity). Secondly, all muons will now precess at the same rate—that is, the spin-effective field will now be homogeneous, although the momentum-effective field (to give focusing) will still be inhomogeneous. This is done by using an electric quadrupole field as well as a magnetic field. Moreover by working at the "magic" energy value $\gamma = 29.3$ (so $\gamma^2 = a^{-1} + 1$), we are unaffected by electric field errors. Finally the relativistical-dilated muon lifetime will be longer.

4. References

[1] Bailey, J., Bartl, W., von Bochmann, G., Brown, R. C. A., Farley, F. J. M., Jöstlein, H., Picasso, E., and Williams, R. W., Phys. Letters **28B**, 287 (1968).

[2] Hague, J. F., Rothberg, J. E., Schenck, A., Williams, D. L., Williams, R. W., Young, K. K., and Crowe, K. M., these Proceedings, and Phys. Rev. Letters **25**, 628 (1970).

[3] deVoe, R., McIntyre, P. M., Magnon, A., Stowell, D. Y., Swanson, R. A., and Telegdi, V. L., these Proceedings.

[4] Lautrup, B., CERN TH-1191 (July 1970), and to be published; and Lautrup, B., Peterman, A., and de Rafael, E., CERN TH-1190.

[5] Brodsky, S. J., and Kinoshita, T., Cornell CLNS-116 (July 1970), submitted to 1970 High-Energy Physics Conference in Kiev.

[6] Drell, S. D., and Pagels, H. R., Phys. Rev. **140**, B397 (1965); and Parsons, R. G., Phys. Rev. **168**, 1562 (1968).

DISCUSSION

D. KLEPPNER: It isn't quite clear to me why the effective field operating on the spin is different from the effective field operating on the orbit? Is there some simple physical explanation?

J. BAILEY: Okay. It's perhaps a bit deceptive. This very much surprised me when I came to look into the equations, the fact that the equations come out like that. Let me see if I can scribble a bit on the bottom here. To your surprise, when you look at the rate at which the spin precesses and the rate at which the orbit precesses, you find that ω_s, the rate at which the spin itself precesses, is of the form $[1+\gamma a]\omega_c$, where ω_c is omega cyclotron. What is in brackets here, this $1+\gamma a$, is half the g factor. That is most surprising because what you're saying is that when gamma is 1, when the body is at rest, half the g factor is 1 plus a. And then you're saying to yourself that under a relativistic transformation the 1 stays and the a part is multiplied by gamma. It's very funny. It isn't that the g factor as a whole gets multiplied by gamma or root gamma or anything, but the two bits of the g factor go up at this different rate, transform each in a different way. It is odd. I don't know why it is. But I must say that the very first time I started to work I rushed off into a corner with a high-precision calculating machine, and I slid the slide along, and sure enough it agreed with the value figured by this equation. It's very surprising. Well, anyway, if you say spin is precessing around at this rate and you subtract the rate that the orbit is precessing around, you get the difference at which the spin is drifting.

In English the word "muon" is very often confused with "moon," and there have been some very stimulating titles of seminars based on this confusion. Well, there is a very close resemblance between muon and moon. We see only one side of the moon. An observer from somewhere like Jupiter would say, "Oh, it's because the moon is spinning on its own axis and also revolving around the earth at exactly the same rate." But if the moon had an anomalous g-factor and these rates were a little bit different to that observer way up there, well, then, sure enough we on the earth would gradually see the other face on the moon. And that is how all these things work, just like the moon. It's very interesting that if you are in a certain magnetic field, it doesn't matter whether you have muons at rest or enormous gamma. They always progress at one-fourth of a megahertz. It's funny. I don't know why. But it works perfectly.

G. R. HENRY: I think that when you showed how you can very cleverly combine the electric and magnetic fields to produce optimum effects in all directions, the effective orbital magnetic field is that magnetic field which would be required all by itself to make a particle move in a circle and—I hope that is correct—and that magnetic field, therefore, is the magnetic field in the rest frame of the laboratory? The magnetic field which you speak of for the spin, isn't that the effective magnetic field in the rest frame of the particle?

J. BAILEY: Sure, it is essentially what you mentioned if you like.

B. M. KINCAID: What kind of numbers did you get? What kind of errors?

J. BAILEY: Oh, well, in the old experiment we had 300 parts per million error on g minus 2 itself, and the new experiment is better by a factor of 20 or so, so we will get an error of about 20 parts per million, we expect.

B. M. KINCAID: Is there an ultimate theoretical limit on your experiment?

J. BAILEY: Well, the statistics, you know. You can only get in a reasonable amount of time before people start to grumble.

An Experiment to Measure g-2 of a Very Low Energy Free Electron

Brian Kincaid, William M. Fairbank, and L. V. Knight

Stanford University, Stanford, California 94305

The experiment to measure the force of gravity on an electron has been modified to measure $g-2$ on very low energy electrons. The apparatus will be discussed and any experimental results will be discussed. In preliminary experiments $g-2$ of the electron has been measured to an accuracy of 30 percent by observing the change in the flight time of ground state electrons in a magnetic field gradient.

Key words: Anomalous moment of the electron; g-factor; low energy electrons.

GRAVITATIONAL CONSTANTS

Recent Developments in the Absolute Measurement of Gravitational Acceleration

A. Sakuma

Bureau International des Poids et Mesures, Sèvres, France

Metrological and geophysical applications of the absolute determination of gravitational acceleration are described and recent work is reviewed. Practical absolute measurements are still restricted to two conventional methods, the pendulum and the free motion experiment. The precision of recent measurements by free motion methods is a few parts in 10^9 ($\sim$1 μgal), two orders of magnitude better than that of absolute determinations prior to 1966. Realization of transportable absolute apparatus is proceeding in several laboratories. The first portable apparatus has recently contributed to the establishment of several new absolute sites (0.1 mgal or better) in North America and also to mutual verification of the accuracy of measurements at two existing absolute sites: NPL in Teddington and BIPM in Sèvres. Agreement to better than 0.1 mgal was obtained between these three absolute determinations. It is therefore believed that there are now no significant systematic errors in the determination of the absolute gravitational acceleration in North America and Europe. Perturbing effects in gravity measurement are analyzed.

Key words: Free motion method; gravitational acceleration.

1. Introduction

Until only a few years ago, a major problem in metrology was how to determine the acceleration due to gravity, "g" in terms of the fundamental units of length (the metre) and time (the second) with an accuracy of 1 mgal (gal = cm/s^2, 1 mgal$\simeq 1\times10^{-6}g$). The details of these efforts have been reviewed by Cook (1965) [1] and others [2].

The principal reasons for this interest were

twofold: First, because of the large scatter in the reported determinations and the incompatibility of the results obtained under different experimental conditions, no previous absolute determination of gravity could be clearly judged accurate to better than 1 mgal. Secondly, local values of gravity, correct to this order, were needed by several laboratories for use with other precise physical measurements whereby g acts as a known transfer constant to express standards of force in terms of the primary standard of mass.

Fortunately, the problem is now essentially solved and there appear to be no further technical difficulties preventing the absolute determination of g within 1 mgal.

This state of affairs has been made possible through the successful completion of several determinations since 1967, all of which required careful preparation and analysis of results over long periods of up to 10 years. A point worth noting is that each of these six independent experiments was based on the principal method for the measurement of g, namely the observation of free motion under gravity, which requires essentially no systematic correction on the observation.

Good agreement has been obtained between these values, and these results have confirmed the validity of some of the prior determinations. On the other hand, they have revealed a great error of about +13.8 mgal in the Potsdam Gravimetric System, according to which the worldwide gravity net has been set up.

Based on these experimental facts, the gravity values of the Potsdam System were revised by −14 mgal in 1967–68 by the International Association of Geodesy [3] and by the International Committee of Weights and Measures [4]. The starting value of the system at Potsdam, originally determined absolutely in 1906 [5], is now taken as 981 260 mgal and no longer as 981 274 mgal. Thus the local values of g at the principal sites on the Earth in the revised Potsdam System are estimated today to be known to an accuracy of 0.5 mgal or better.

An uncertainty of 1 mgal in gravity affects other precise physical measurements in the following ways: 0.5 ppm in the absolute ampere, 1 ppm in the absolute volt, force and pressure, and a 0.000 03 K error in the realization of the fixed temperature at the boiling point of water; therefore the contributed error is at present much less than the other errors arising from perturbing elements in these measurements.

Thus, the aim of the dominant use of gravity in metrology has now been attained.

2. Absolute Gravimetry in Geophysics

Almost all gravity measurements in geophysical studies simply require the measurement of small differences in gravity from place to place or the measurement at one place of very small variations

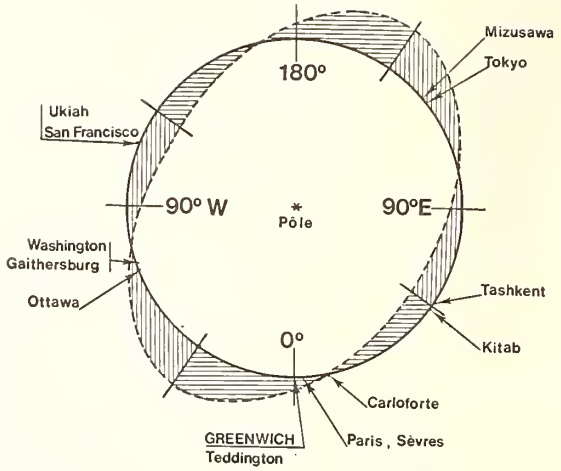


FIGURE 1. *Probable oscillation of geopotential of ∼19 years period, predicted by the polar motion observations (1933–65): Okuda [6].*
Circle represents the Earth's section at latitude N-39°08′.

(∼0.1 mgal) due to the effects of the Earth's tides. Consequently the techniques of absolute determination had been considered quite inadequate for these purposes because of the lack of sensitivity and the difficulties of transportation. However the situation in absolute gravity is changing rapidly and recent experiments confirm that the value of g can be determined absolutely with an accuracy of a few microgals ($\sim 10^{-9}g$) and also that gravity differences can be measured with a few hundredths milligal accuracy ($\sim 10^{-8}g$) by portable absolute gravimeters. These experimental advances are very important because only an absolute apparatus of such high sensitivity permits the study of the long-term variation of gravity, the so-called secular variation of g, resulting from a number of important geophysical and cosmological phenomena: crustal movements, evolutions of the form and inner structure of the earth, secular change of dynamic constants in the solar system, eventually the variation of the Newtonian constant of gravitation G, and so on. Moreover, portable absolute gravimeters allow the covering of many fundamental points on the Earth with greater accuracy than values in the Potsdam System which have been set up with relative gravimetric measurements, and will in future render this system obsolete.

An interesting phenomenon in the secular variation of the geopotential, or the vertical line of the Earth, was recently predicted by Okuda [6]. This hypothesis is derived from the periodical variation of the discrepancy in data on polar motion (1933–65) obtained by five different observatories (Carloforte, Kitab, Mizusawa, Ukiah, and Gaithersburg) situated at latitude N-39°08′, and concluded that the geopotential or the vertical line is oscillating with a period of about 19 years as shown in figure 1; Europe (Carloforte) is on the node, and the Far East (Mizusawa) and the West Coast of North America (Ukiah) are near the loops of this oscilla-

tion. Therefore the secular variation of gravity will be observed much more clearly at places near the loops and it is expected that the variation will attain $10^{-7}\sim10^{-8}g$ during the period of 19 years.

Because of its complete freedom from instrumental drifts, absolute gravity measurement is considered the best way to verify this theory and a project is now in progress at the International Latitude Observatory, Mizusawa, to install an absolute gravimeter of high sensitivity dedicated to the observation of gravity variation as well as to the establishment of one of the most reliable absolute stations in the world gravity net.

In this way, the modern absolute determination of gravity is now finding new applications in geophysical studies, and such applications also provide new interests for metrology as a number of technical refinements and improvements are required in length measurements, stabilization of laser wavelength, high-speed chronometry, vibration and temperature controls, precision mechanics, and so on.

3. Methods for Absolute Determination

In the past, when the precision of gravity measurements was not as high as today, one could imagine several auxiliary methods devised especially for the facility of time measurements, such as Galileo's tilted plane, Atwood's machine, simple pendulums, rotation of liquid surface method, and so on. However, referring to the obtainable precision by modern determinations, $10^{-7}\sim10^{-9}g$, those methods are no longer of interest so far as precision is concerned, and only two classes of methods are worth noting here, the free motion and the reversible pendulum.

3.1. Free Motion Under Gravity

In spite of its simplicity, this well-known principle has been noteworthy only since the work of Ch. Volet [7] in 1946 when high-speed timing techniques became available. Practically, this principle is divided in two classes, symmetrical free fall and simple free fall; the former is more advantageous metrologically than the latter.

a. Symmetrieal Free Motion

This method, also proposed by Ch. Volet [8] in 1947, is shown in figure 2a. An object is projected vertically upwards crossing two defined horizontal stations, S_1, S_2, whose vertical separation H is known. Two independent time intervals corresponding to the upward and downward passages across each station T_1, T_2 are measured and the value of g is obtained by eq (1). Assuming that the uncertainties of time measurements, ΔT_1, ΔT_2, are inversely proportional to the velocity of the object and letting $H=1$ m, $\Delta H=5$ nm, $T_1=0.2$ s, $\Delta T_1=5$ ns, $T_2=0.93$ s, $\Delta T_2=1$ ns, the standard deviation of g is calculated to be $\Delta g=\pm0.007$ mgal; therefore the experiment makes an optimum use of the available accuracy of measurements of distance and time.

Figure 2-a

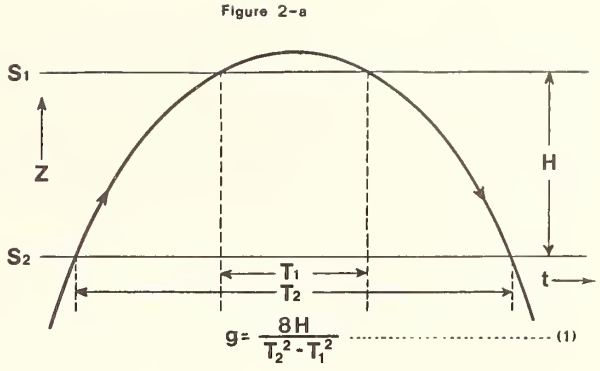


Figure 2-b

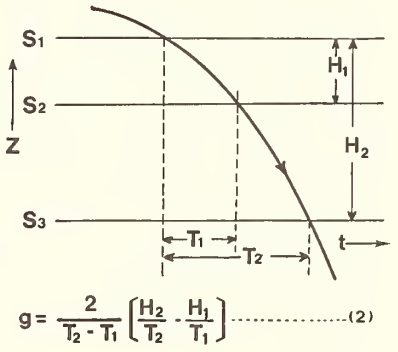


FIGURE 2.

There are two important advantages of the method. Firstly the time intervals are defined at each station by two signals of equal form and sharpness due to the equal velocity of the object passing each station in upward and downward sense. Therefore, timing errors coming from a time constant (limit of bandwidth) of chronographs are cancelled out. Secondly, the value of g obtained by this method is free from the first-order resistance of residual air (proportional to the velocity of the moving body) in a vacuum vessel, due to the fact that during the upward motion the resistance acts in the downward direction and during the downward motion the direction of the resistance is reversed in such a way that the time intervals T_1, T_2 are independent of the resistance. Similarly, another form of resistance caused by eddy currents due to inhomogeneity of the magnetic field along the trajectory is also cancelled out.

On the other hand, a practical difficulty of the experiment is that mechanical shocks caused by the launching of the object are liable to disturb the observation, and so the catapult must be designated so that the shocks and the rotation of the projected body are as small as possible, and the form of the body is to be chosen in a way that its rotation does not produce additional errors in this experiment.

b. Free Fall Method

Since the initial velocity and position of a free falling object are difficult to determine precisely in

this method, figure (2b), the two successive times of passage of the object, T_1, T_2, must be measured through three horizontal stations S_1, S_2, S_3 for which the vertical separations H_1, H_2 are known. The acceleration g is obtained by eq (2), figure 2b. Conversely the value of g can also be determined by measuring the two successive distances of fall H_1, H_2 travelled by the object during two given time intervals T_1, T_2.

The standard deviation calculated from eq (2) may have no important meaning because of strong correlations between two successive time intervals T_1, T_2. Nevertheless, to permit a comparison of the order of uncertainty of the method with the symmetrical free motion under equivalent conditions, let $H_1 = 0.33$ m, $\Delta H_1 = 1.6$ nm, $H_2 = 1$ m, $\Delta H_2 = 5$ nm, $T_1 = 0.18$ s, $\Delta T_1 = 3.7$ ns, $T_2 = 0.36$ s, and $\Delta T_2 = 3.5$ ns; then the standard deviation of g obtained is $\Delta g = 0.09$ mgal. This result is over 10 times less sensitive than that of the symmetrical free motion method. The disadvantages of the experiment come (i) from the large variation of the velocity of the falling object, so that almost all of the timing errors are caused at the uppermost station S_1 when the velocity is lowest, and (ii) from the shorter time intervals to be measured T_1, T_2 compared with T_2 of the symmetrical free motion method. Care must be taken against the resistance of residual air and the phase shifts of signals due to a time constant of the chronographs. Practical advantages of the methods are the ease of dropping and the smaller mechanical shocks caused by release of the body than in the case of symmetrical free motion.

3.2. Reversible Pendulum

The well-known principle of the reversible pendulum is that, if a body oscillates as a pendulum about two centers with the same period, the distance of the centers is equal to the length of the simple pendulum with that period. This is the only constrained-motion experiment under a gravity field in which an accuracy of a few tenths of mgal ($\sim 10^{-7} g$) may be obtained under ultimate experimental conditions. The limit in accuracy comes from its principle: the length must be determined in the oscillating state under a strong constraint between the supports (knife-edge and bearing plane) of the pendulum. The localized pressure between them attains ~ 1000 N/mm^2 and so considerable plastic deformation and wear of knife-edge and plane may take place. Due to the low velocity of the swing, e.g., a few mm/s, corresponding to that in free fall after only a few μm drop from rest, a large number of periods must be integrated, say over one hour, for precise period determinations. This long observation had been an important advantage of timing in the past; all three important absolute determinations before 1940 [5, 9, 10] had been performed by this principle. However this is no longer an advantage today.

Thus the reversible pendulum is employed less often in modern absolute determinations, and the future of the experiment seems to lie in studies of the behaviour of the pendulum itself under a known gravity field. Improvements of the accuracy and eliminations of systematic errors of the pendulum may then be made to permit the use of the pendulum as an auxiliary instrument for high-precision gravity measurements.

4. Actual Status of Gravity Measurements

Six absolute determinations have been completed since 1967 (table 1) and eight other experiments are in progress.

4.1. Completed Determinations Since 1967

a. National Bureau of Standards (NBS), Gaithersburg

This is a typical example of the free fall experiment in which the two distances are defined on a falling body itself and time intervals of free fall corresponding to the distances are measured at one fixed point. The falling object is made of a fused silica rod of 1 m length having three optical slits 30 cm and 1 m distant. The rod falls in a vacuum enclosure which itself falls approximately freely in the atmosphere guided by two vertical columns. Because of a very low relative velocity between the rod and the enclosure during the fall, the resistance of residual air on the falling rod is considerably reduced and the correction is calculated as less than 0.03 mgal at a pressure of 4 μbar. Timing signals are obtained by a photomultiplier when the slits of the falling rod intercept a light beam. In this experiment the light beam incident to the photomultiplier passes three different windows fixed on the vacuum enclosure. Small tilts of the windows or deformation of the enclosure therefore displace the light beam and timing errors take place; this effect was the first difficulty to be solved. A final result was obtained with a standard deviation of 0.3 mgal, which is estimated as the best among the results derived from the line standard of length.

b. National Physical Laboratory (NPL), Teddington

This is the first symmetrical free motion experiment which proved experimentally the advantages of this method. A glass ball of 2.5-cm-diam, acting as a lens, is used as the projected object, in order to eliminate the effect of rotation. Timing signals whose halfwidths correspond to about 3.5 μm displacement of the ball are obtained when the ball passes between two horizontal slits and focuses one on the other. Two such slit pairs mounted on optically worked glass blocks are placed 1 m above the other and the distance between them is determined interferometrically with reference to a Fabry-Perotetalon of 20 cm enclosed in a constant-pressure case. Electrical charging of the ball at high vacuum had been the main difficulty in the preliminary work and the final determination was made at a relatively low vacuum of 10 μbar at which the charging is negligible. Air resistance has been shown to be unimportant in this

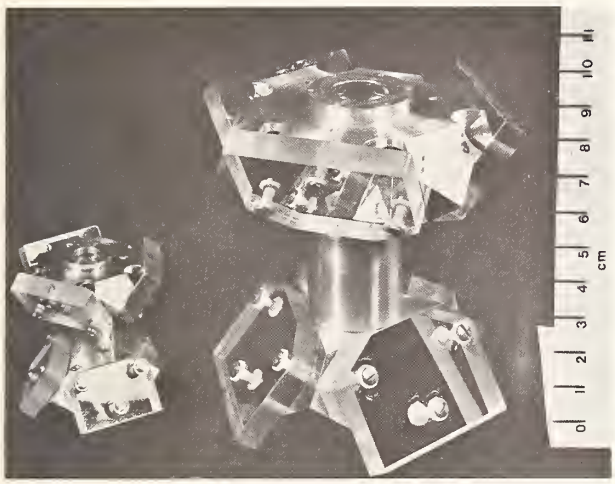


FIGURE 3. *Cube-corner mirrors for free motion gravity measurement: left, 70 g for portable apparatus, right, 430 g for fixed apparatus.*

experiment and the value of g is not in error by more than 1 mgal when the ball is projected in a low vacuum of 0.7 mbar. The final results had a standard deviation of 0.13 mgal of which the major part was due to microseismic ground motion.

c. Bureau International des Poids et Mesures (BIPM), Sèvres

This is another symmetrical free motion experiment in which a cube-corner mirror, 10 cm high and weighing 430 g (fig. 3), which forms one mirror in the vertical beam of a Michelson interferometer, is used as the projectile (fig. 4a). The two stations are installed in the trajectory as the conjugate points of two mirrors forming an end standard of 80 cm length made of fused silica, placed in the horizontal beam of the interferometer. Timing signals by white light fringes with a halfwidth corresponding to a 0.06 μm displacement of the cube-corner mirror, are obtained by a photomultiplier when the optical path difference between the horizontal and the vertical is null. The frequency of the fringes is very high, about 30 MHz; in addition the effective surface of mirrors is made very small, 0.2 cm^2, in order to assure highly plane surfaces on the mirrors. Thus an ordinary white light source is not sufficient to get a good S/N ratio for the fringes, and so a xenon flash lamp is synchronously triggered with each passage of the cube-corner mirror at the two stations in upward and downward motions [16]. Care must be taken to eliminate the effect of microseismic disturbances whose vertical component enters, in the value of gravity, as a first-order error, and therefore the vacuum enclosure containing the interferometer is mounted on a table stabilized by a servocontrol system (fig. 5).

In addition the residual vibration of the interferometer is recorded, during the flight of the object, by a long-period vertical seismometer (30 s) in a tight pressure case with magnetic shielding. The

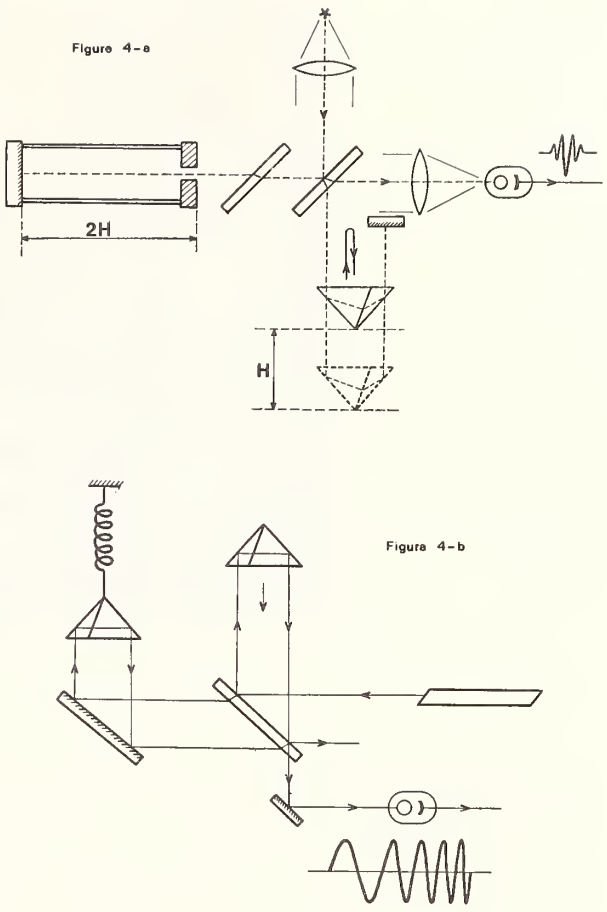


FIGURE 4.

length of the reference end standard is measured periodically, in its place in vacuum, by a krypton 86 lamp cooled at 58 K with an optical path difference of 80 cm. The length is confirmed to stay constant at 20 °C to 1×10^{-9} over one year referring to the Kr 86 lamp, but a difficulty arises from the fact that the meter is not defined with this accuracy. Therefore, a most probable wavelength of Kr 86 ($\lambda = 605\ 780\ 199.8$ fm) is used for a length determination which seems concordant by a few parts in 10^{-9} with the meter definition. The perturbation of gravity due to the Earth's tides ($\sim$0.1 mgal) can be measured by this apparatus (fig. 6). After three years of observation of the value of gravity at the reference point, Sèvres A 2, the secular variation of gravity and the seasonal variation of gravity are confirmed as not exceeding 0.02 mgal. Scatter of single measurement is influenced by meteorological conditions but as a mean 0.005 mgal is obtained.

d. Wesleyan University, Middletown

This is the portable absolute apparatus ($\sim$800 kg) with which the first absolute gravity liaisons between Europe and the North American continent were carried out in 1968. In this free fall experiment,

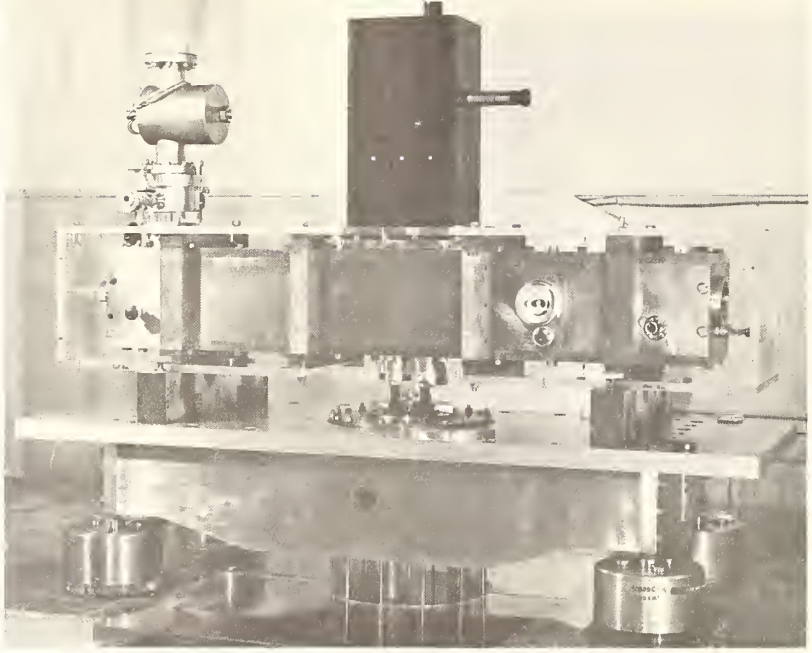


FIGURE 5. *A prototype of absolute gravimeter: precision* $\sim 10^{-9}$ *g.*
Interferometer in vacuum vessel is placed on a stabilized table. (BIPM).

the fringe-counting method with a stabilized Hc–Ne laser is used for length measurements as well as for time measurements (fig. 4b). A cube-corner prism ($\sim$300 g), forming one reflector in the vertical beams of a double-beam Michelson interferometer, falls freely about 1 m in a relatively high vacuum ($\sim$1 nbar). Time measurements are made between three moments when the high-speed interference fringes (3.5 to 15 MHz) detected by a photomultiplier cross the zero phase positions and simultaneously two whole numbers of fringes corresponding to these time intervals are counted.

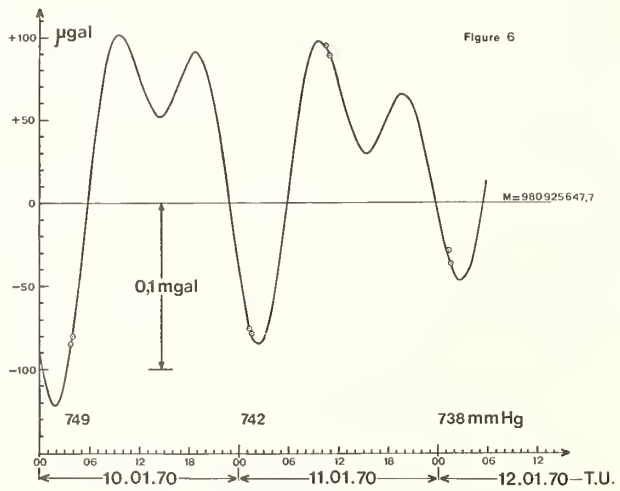


FIGURE 6. *Comparison of eight absolute determinations of g to a theoretical curve of the Earth tide; calculated on 378 waves taking account of the local phase shifts of the waves.*

Electronic circuits are carefully checked to avoid errors coming from the phase shifts. Owing to the high sensitivity of the interferometric detection, an accuracy in the time measurements of 1 ns is obtained over 0.36 s. Microseismic disturbances are much reduced by suspending a reference cube-corner prism in the horizontal beams from a vertical seismometer by means of a 45° inclined mirror.

Small corrections are applied to the results; one is a correction due to the Doppler shift of the wavelength of the laser by the dropping cube-corner prism, 0.03 mgal in this case, the other is a correction of air resistance and is of the order of 0.01 mgal in a pressure of $\sim$1 nbar (1×10^{-6} mm Hg).

One difficulty encountered in this experiment is a weak magnetic interaction between the seismometer and a solenoid which releases the falling prism. Two different mean values of gravity were obtained by currents in opposite directions in the solenoid, and sometimes the difference reached about 0.1 mgal. However the mean of these two values is considered to be very near the true value of gravity. The results with this portable apparatus (table 1) at two absolute sites in Europe, NPL Teddington and BIPM Sèvres, were in good agreement with the others. At the same time, the apparatus demonstrated the ability of such an absolute gravimeter to establish a more accurate absolute world gravity net than the Potsdam System.

e. National Research Laboratory of Metrology (NRLM), Kakioka

The method used in this experiment is a free fall of a standard scale of 1 m length, traced every 1 cm. About 100 intermediate passages of the falling scale

TABLE 1. *Results of absolute determination of gravity since 1965*

Site of determination	Author and date	Value implied at site indicated (mgal)		Difference from revised Potsdam system
NBS, NBS-2 Gaithersburg	Tate 1966 [11]	Commerce Base Washington	980 104.8 ±0.5	+0.8 ±0.5
NBS., NBS-3 Gaithersburg	Hammond[a] and Faller 1969 [12]	Commerce Base Washington	980 104.24 ±0.08	+0.24 ±0.08
NPL, Bushy House Teddington	Cook 1967 [13] revised 1969[b]	British Fundamental Station	981 181.81 ±0.13	+0.34 ±0.14
NPL, Bushy House Teddington	Hammond[a] and Faller 1969 [12]	British Fundamental Station	981 181.865 ±0.060	+0.40 ±0.08
BIPM, Sèvres-A2, Sèvres	Sakuma 1970 [14]	Sèvres A	980 925.949$_0$[c] ±0.0054	+0.22 ±0.07
BIPM, Sèvres-A, Sèvres	Hammond[a] and Faller 1969 [12]	Sèvres A	980 925.965 ±0.050	+0.23 ±0.09
NRLM, Kakioka	Kitsunezaki et al. 1969 [15]	Kakioka Gravity Station	979 964.8 ±2.0	−1.3 ±2.0
PTB, Braunschweig	German 1970[d]	Apparatus Site	981 252.3 ±2.0	+1.0 ±2.0

[a] Wesleyan University.
[b] Cook, A. H., and Hammond, J. A., Metrologia 5, No. 4, 141–142 (Oct. 1969).
[c] Value in August and September 1969.
[d] Private Communication.

are detected by a photoelectric microscope and the waveforms displayed on an oscilloscope with the time scales are photographed on a running film. Particular care is necessary to avoid rotations of the falling scale, thus assuring good sharpness of the images of the traces. Two kinds of scales were used, one of invar and another of fused silica; no systematic errors were found between the results obtained by each of the scales. One major part of the scatter of gravity values, $\sigma = \pm 2$ mgal, was due to oscillations of the photoelectric microscope which was fixed directly on a sidewall of a vacuum vessel in which the scale falls.

f. Physikalisch-Technische Bundesanstalt (PTB), Braunschweig

This is another free fall experiment using high-speed photography. The falling body is a fused silica rod 2 m long coated with photographic emulsion. A number of images of an optical slit illuminated by short-duration light pulses, 0.1 μs, 130 Hz, are photographed on the falling bar. Images from a large number of drops can be recorded on the same bar by dropping it with different azimuths. The main uncertainty comes from the difficulty of length determinations which were made between non-uniform images caused by the rapid variation of the falling velocity. The standard deviation of the determination of 2 mgal and the value in table 1 is a provisional one, but this has the character of a final value.

4.2. Absolute Determinations in Progress

a. National Standards Laboratory (NSL), Chippendale

This is a symmetrical free motion experiment on much the same line as that at BIPM (fig. 4a), but instead of a flash lamp, a bright arc lamp is used as white light source and a silicon photodetector with high efficiency in the near infrared region is used. Thus white light fringes of about 0.9 μm mean wavelength are obtained in the infrared. In preliminary experiments, some high-frequency vibration was transmitted to the measuring system by the movement of the catapult piston and the mechanism was redesigned. An accuracy of 0.1 mgal is currently being obtained and the experiment is noted as the first absolute determination in the Southern Hemisphere.

b. Thompson-CSF, Paris

This is the first trial at industrialization of the portable absolute gravimeter by the free fall method. The apparatus is very similar to that of the Wesleyan University with its laser fringe-counting system (fig. 4b), but is made much more compact, having a total mass of 130 kg. A preliminary result [17] was about 0.15 mgal less than the value obtained by a gravimetric tie with the BIPM and this difference is considered due to the air resistance by a relatively low vacuum of 40 nbar. This first apparatus is expected to be completed in 1971.

c. Istituto di Metrologia "G. Colonnetti," Torino

This is also a portable absolute gravimeter with the laser fringe-counting system and is being done in collaboration with the BIPM. Symmetrical free motion is employed in this apparatus, but the time measurement is made at only one station, S_2 in fig. 2a, $(T_1=0)$, and the height from the station to the apex of the trajectory H is determined simultaneously as a function of the laser wavelength. High vacuum is not required in this apparatus and experiments will be made in a pressure of about 15 μbar.

d. Deutsches Amt für Messwesen und Warenprüfung, Berlin

This is a free fall system but instead of a long falling object being dropped, a small body provided with slits drops through three light beams at distances of about 1 m.

e. Air Force Cambridge Research Laboratories, Bedford

This is a reversible pendulum determination; the pendulum, simple length of 7.5 cm, is made with a rectangular frame composed of four blocks of fused silica in optical contact. The length of the pendulum can be determined interferometrically and a repeatability of 0.3 mgal is obtainable.

f. Geodetic Institute, Potsdam

This is also a reversible pendulum experiment, in which two pendula are swung together with opposite phase to eliminate flexure of the support and the effect of ground motion.

g. Finnish Geodetic Institute, Helsinki

The aim of this experiment is to determine the value of gravity by a long pendulum consisting of a bob hung on a fine wire. As the first step of the experiment, pendula of 4 m and 8 m are used and a number of perturbing effects are being studied. A 3 mgal reproducibility is presently obtained.

h. University of Buenos Aires

This is the succession of the reversible pendulum experiment by Heyl and Cook (1936) at the NBS. Reconstruction and adjustments of the apparatus are in progress.

4.3. Remarks on the Experiments

Between the results obtained by each of the six laboratories (table 1) there are differences in precision of about two orders of magnitude. These differences are due to the choices made of the three following experimental factors; (i) the symmetrical free motion method or the free fall method, (ii) length determination by interferometric devices or by line standards, (iii) vibration is controlled or not controlled. Of course the former of the three are advantageous; in fact the advantages of the interferometric devices are fourfold: (i) length determination can be made in optimum conditions, (ii) moving body (mirror or prism) can be made both small and rigid, (iii) very sharp signals by interference are obtained for precise time measurements, (iv) no material stations are required in the trajectory of the moving body, so the vibration controls and the temperature controls in the area of the trajectory are not important.

Among the completed experiments, only one example (4.1.C., BIPM) has adopted all of the three advantages and the result is clearly superior. An important feature of this experiment is that the precision obtainable is principally unlimited and no systematic correction is required. Therefore precise gravity measurements in laboratories in the future will tend to adopt this method with several technical refinements for researches of $10^{-9}g$ and it seems probable that such apparatus will be installed at several sites on the Earth.

Another important feature of modern gravity measurement is the realization of portable gravimeters whose utility has been proved by Wesleyan University and this kind of apparatus will be used often, in the future, for geodetic purposes; moreover the apparatus can be constructed in industry, just as in the case of the spring relative gravimeters.

In this way, it is an inevitable evolution in science that the place of the classical absolute determination of gravity with 1 mgal accuracy is becoming less important and such experiments will be undertaken less often, in the future, except for some special purposes in physical experiments other than the gravity measurement itself.

5. Perturbing Effects on Gravity Measurement

Several perturbing effects on precise gravity measurement are summarized and some experimental results are described below.

5.1. Horizontal Velocity (Eötvös Effect)

As is well known, centrifugal acceleration by the rotation of the Earth plays a role in gravity. Therefore, if a horizontal velocity in the east–west direction is given to the falling object, the observed value of gravity is changed. This effect is about $1\times10^{-8}g$ at latitude 50° at a velocity of 1 mm/s. Experimentally, it is possible to reduce the velocity up to

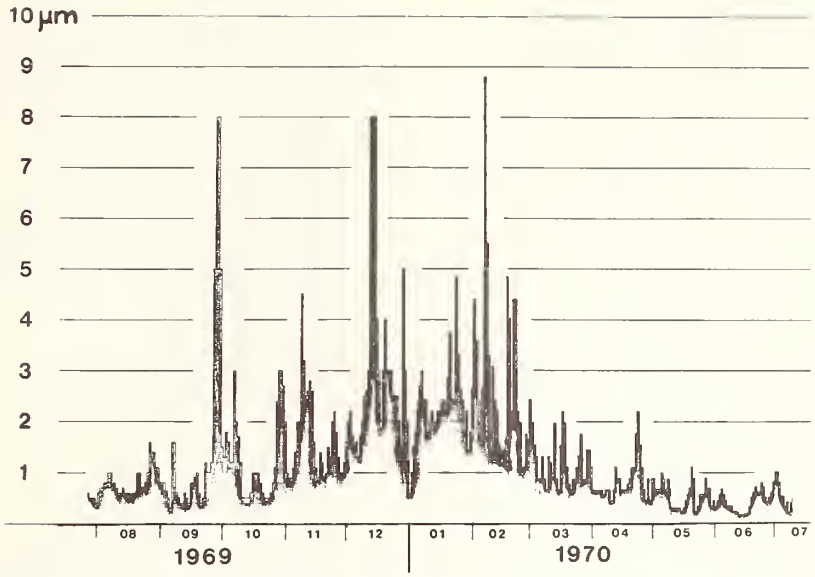


FIGURE 7. *Mean ground noise (vertical, peak to peak amplitude) due to oceanic microseisms, observed at Sèvres.*

$\sim$0.2 mm/s ($2\times10^{-9}g$) by careful adjustment of the catapult, but if a better reproducibility is required it is necessary to measure this velocity in each experiment.

5.2. Oceanic Microseisms

Ground motion contains, in general, two separate spectra of microseisms. One is industrial ($\sim$5–50 Hz) and its amplitude varies very much from place to place; the other is oceanic (0.15–0.4 Hz), whose amplitude changes considerably under meteorological conditions, as seen in an example in figure 7. If this effect is not compensated, the disturbance of gravity attains $\pm$0.16 mgal by a microseism of 0.2 Hz, 2 μm peak-to-peak amplitude. Moreover these microseisms are very difficult to measure accurately because of their low frequencies and generally several percent of the acceleration is transmitted to a long-period seismometer of 30 s. One practical solution to eliminate this effect is to compensate roughly for the microseisms, for example, to one tenth of their initial amplitudes, by piezoelectric elements and the residual of the microseisms is measured by a seismometer. By this procedure, acceleration due to microseisms can be reduced over 100 times.

5.3. Barometric Pressure

Variation of barometric pressure influences gravity measurement in several ways: (i) a variation of local pressure changes the gravitational attraction of upward direction due to the mass of air. This effect is calculated to be about 0.006 mgal/cm Hg. But this pressure variation also causes deformations of the ground so that the effective gravity variation is reduced from the calculated value [18]; moreover this effect is very much influenced by local geological structure of the ground and accurate measurement is very difficult, at present, due to instrumental deformations of relative gravimeters by pressure changes. Absolute gravimeters of high sensitivity will make it possible to study this phenomenon more clearly in the future. (ii) Periodical variation of barometric pressure with periods of a few minutes produces deformations of vacuum vessels and changes the air buoyancy; these effects are clearly observable on a windy day. Therefore, pressure variation must be observed by a sensitive barometer in order to measure gravity at the moment when the pressure is constant. It is very important that the microseismometer be installed in a tight pressure case to avoid the variation of air buoyancy. (iii) Opening of doors and acoustic vibrations are also important perturbing effects in gravity measurement, and these possible effects must be avoided at the moment of the measurement of gravity.

5.4. Terrestrial Magnetism

The magnetic field can be changed by several percent because of traffic and industrial activity, and this fluctuation is liable to excite the spring of sensitive microseismometers as seen in the experiment of 4.1.d. Therefore the seismometers must be protected against this effect by magnetic shields.

5.5. Mass Distribution

As is well known, a mass of 1 ton creates a gravitational acceleration of $6.8\times10^{-9}g$ (0.006 7 mgal) at a place distant 1 m from it. Thus, care must be taken to avoid mass displacement in the vicinity of the gravimeter, especially when the long-term stability of gravity is studied.

5.6. Teleseisms

Remote earthquakes, several hundred times per year, disturb a gravity measurement considerably as the Earth is moved 1 mm or more, at frequencies of 0.1 to 0.01 Hz, by earthquakes of large magnitude, which occur perhaps several times a year. Therefore, it is very desirable to know if such a disturbance is occurring during the gravity measurement.

6. Conclusion

Recent developments in technology have made it possible to determine the absolute value of gravity with an accuracy up to $10^{-9}g$ after long and patient efforts by scientists of the world over. These techniques have contributed to metrological applications and to the revision of the Potsdam Gravimetric System. The modern techniques of absolute gravimetry are beginning to provide several new means for geophysical studies by which new knowledge will be obtained on the Earth and on the Solar System.

7. References

[1] Cook, A. H., Metrologia **1**, No. 3, 84–114 (1965).

[2] Huntoon, R. D., and McNish, A. G., Del Nuovo Cimento Suppl. **6**, Serie X, 146–184 (1957).

[3] Bulletin Geodésique No. 86, 383 (1967).

[4] Procès-Verbaux C.I.P.M., 2d séries, Tome 36, 57th session, p. 27 (1968).

[5] Kühnen, F., and Furtwängler, P., Veröff. Königl. Preuss. Geod. Inst. **27** (1906).

[6] Okuda, T., On a new problem in the variation of latitude the 19 years periodicity of the local z term (In Japanese), Astron. J. Japan **61**, No. 3, 72–76 (1968).

[7] Volet, Ch., C.R. Acad. Sci. (Paris) **222**, 373 (1946).

[8] Volet, Ch., C.R. Acad. Sci. (Paris) **224**, 1815–16 (1947).

[9] Heyl, P. R., and Cook, G. S., The value of gravity at Washington, J. Res. Nat. Bur. Stand. (U.S.) **17**, 805–839 (1936).

[10] Clark, J. S., Phil. Trans. Roy. Soc., London, Ser. **A238**, 65 (1939).

[11] Tate, D. R., J. Res. Nat. Bur. Stand. (U.S.) **72C** (Eng. and Instr.), No. 1, 1–20 (1968).

[12] Hammond, J. A., JILA Report No. 103, Univ. of Colorado (1970).

[13] Cook, A. H., Phil. Trans. Roy. Soc. London **261**, 27, 211–252 (1967).

[14] Procès-Verbaux C.I.P.M., 2d Séries, Tome 38, 59th session, 1970 (in press).

[15] Report presented to the International Conference on Precision Measurement and Fundamental Constants, Gaithersburg, Aug. 1970.

[16] Sakuma, A., Bull. Géod., No. 69, 249–260 (1963).

[17] Lacombat, M., and Pircher, G., Mesures, Regulation, Automatisme **34**, No. 6, 63–65 (June 1969).

[18] Lecolazet, R., and Steinmetz, L., C.R. Acad. Sci. (Paris) **263**, 716 (1966).

A Laser-Interferometer System for the Absolute Determination of the Acceleration Due to Gravity

James A. Hammond*

University of Colorado, Boulder, Colo. 80302

James E. Faller

Wesleyan University, Middletown, Conn. 06457

A new and portable instrument for making an absolute determination of the acceleration due to gravity has been built. The design and construction of this apparatus together with the results obtained at eight different sites are discussed. The instrument consists of an optical interferometer illuminated by light from a stabilized He–Ne laser in which one of the mirrors, a corner cube reflector, freely falls a total distance of about one meter. The drop-to-drop scatter is less than 1 part in 10^7, and the uncertainties (70% confidence intervals) of the results obtained at most sites are less than ± 5 parts in 10^8.

Key words: Absolute gravity; acceleration of gravity; g; geodesy; gravimeter; world gravity network.

We describe here a new and portable system for the measurement of the absolute acceleration of gravity. Results are given for measurements made at eight different sites. Since the instrument concept has already been described in the literature [1, 2], it will be only briefly described here. A complete description of the instrument and the measurements that have been made with it is available [3].

The instrument is basically a Michelson interferometer which uses corner-cube reflecting elements, one of which falls freely. Corner cubes are used as the reflectors and the dropped object is designed so that its center of mass coincides with the cube's optical center, thus eliminating any optical path change as a result of the rotation of the dropped object during free fall. The interferometer is illuminated by a stabilized He–Ne laser. Figure 1 shows schematically the optical system. The interferometer is aligned by appropriate manipulations of the adjustable mirror and the laser. The light beam in the free fall arm is autocollimated against a mercury pool and adjusted to be vertical so that the interference fringe count corresponds precisely to the distance fallen in terms of the true wavelength of the laser light. To accomplish this, the light in the reference arm is blocked, the mercury pool uncovered and the returning pinhole image observed and centered on the entrance aperture making use of a small beam splitter and the microscope shown in the diagram.

To determine g, we measure the distance the object falls during a known time interval in terms of the number of interference fringes which are detected by a photomultiplier. In practice two intervals are used since the initial velocity is unknown. The expression for the measured value of the acceleration of gravity is:

$$g = \frac{\lambda(M_2 - M_1 \cdot T_2/T_1)}{T_2^2 - T_1 T_2}$$

where λ is the wavelength of the light, and M_2 and M_1 are the number of interference fringes which occur in the time intervals T_2 and T_1, respectively. The intervals T_2 and T_1 must begin at the same time; we usually operated with T_2 set at twice T_1.

The arrangement of the important elements in the mechanical construction of the apparatus is shown in figure 2. The apparatus as it is shown has an overall height of just under nine feet with the measurement being made over a path of about one meter length inside the vacuum chamber. The seismometer shown in figure 2 provides an inertial system for supporting the reference corner-cube to minimize the effects on the measurement of natural and man-made seismic disturbances as well as any shock accompanying the release of the dropped object.

Figure 3 is a schematic diagram of the electronics used to make the actual measurement. The time standard is a 5 MHz precision oscillator which is intermittently compared with WWVB, the National Bureau of Standards 60 kHz standard frequency transmission. This provides a reference which is at all times within 5 parts in 10^{10} of the international time standard. The adjustable time delay generator

* Present address: National Bureau of Standards, Washington, D. C. 20234.

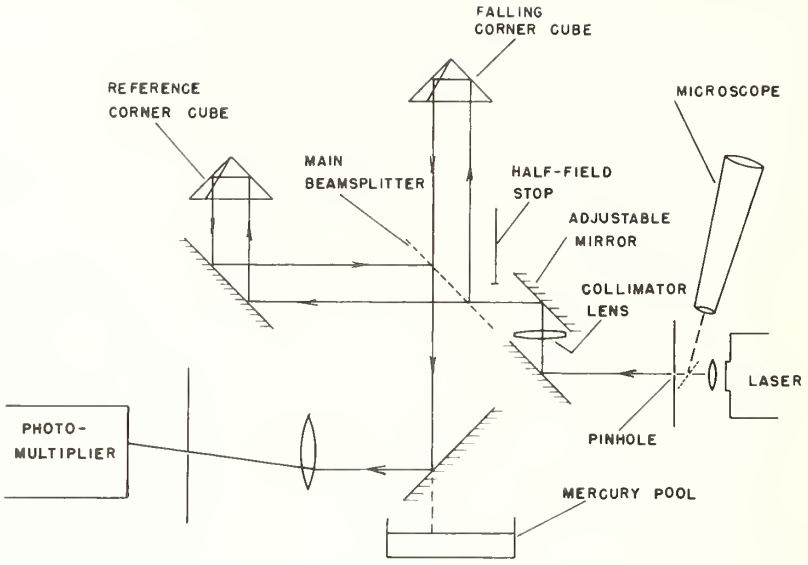


FIGURE 1. *Diagram of optics.*

produces three pulses to start and stop the two channels of the counter at the appropriate times. The two intervals, usually 0.18 s and 0.36 s long, are begun after an initial time interval of about 0.1 s.

The photomultiplier signal is converted to a series of 5 ns (full width half max.) pulses by a zero-crossing descriminator which in turn provides the inputs to a dual counter and also to a time-to-amplitude converter (TAC). The TAC measures fractional fringes by measuring the small time intervals between the counter's start and stop signals and the following fringe pulse. In practice we applied this as a correction to the preset time intervals so that we effectively measured the time intervals corresponding to the integral number of interference fringes as recorded by the two counters. The data necessary to determine the value of gravity are the two fringe counts, M_1 and M_2, the two preset time intervals set by the time delay generator, and the small timing corrections to these intervals as measured by the time-to-amplitude converter.

The apparatus cycles automatically once started and makes 100 drops per hour. The data as acquired in 50 drop sets is averaged and then that average is corrected for the effect of the tidal acceleration of the moon and the sun. This results in the measured

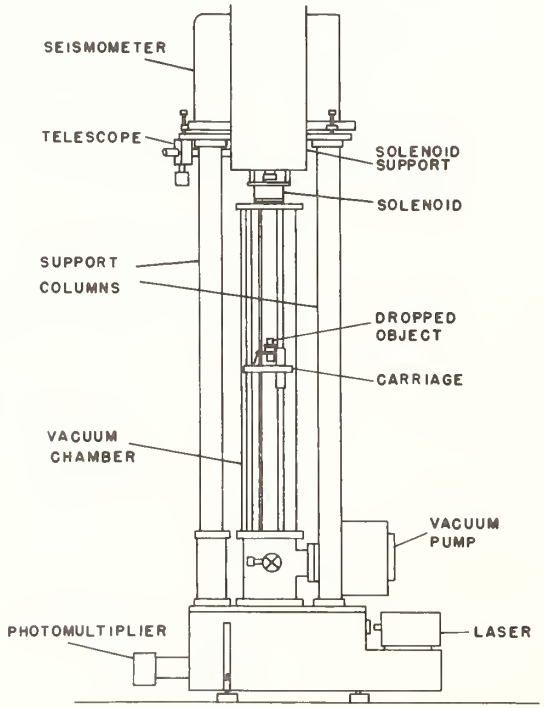


FIGURE 2. *Diagram of the Apparatus.*

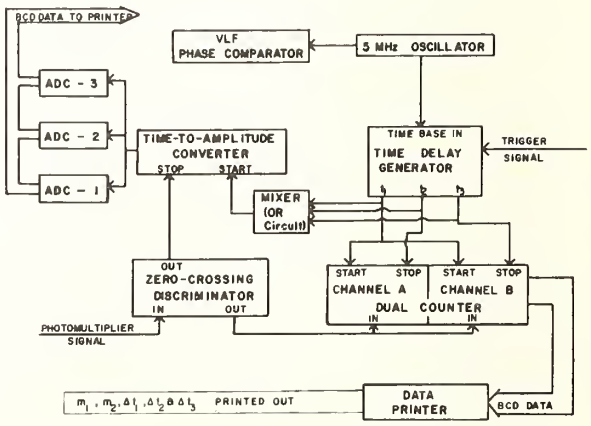


FIGURE 3. *Schematic diagram of the Electronics.*

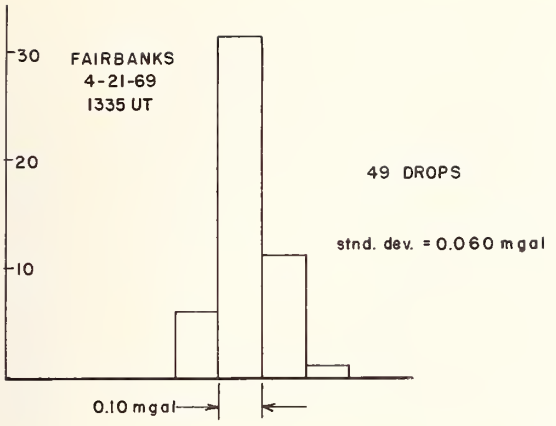


FIGURE 4. *Histogram of data with inertial suspension of reference corner-cube.*

value of g for that half-hour period. The tidal correction is calculated from the formulas presented by Longman [4] for the tidal acceleration using a gravimetric factor of 1.2 to allow for the earth tides. At any particular site a number of these 50 point data sets are taken and averaged. This results in the measured value of g assigned to that site.

Figure 4 is a fairly typical histogram of a set of data; the standard deviation of a set of data is usually about 0.09 mgal.[1] This may be compared with figure 5 which is a histogram of data taken immediately after that in figure 4 but with the inertial suspension system locked so that the reference corner-cube is firmly attached to the supporting structure and hence to the earth. This is representative of the degree of improvement obtained with the inertial suspension at sites where the seismic background was moderate as it is in this part of Alaska. The severest seismic disturbances occurred in Bogota, Columbia where the data scatter was not substantially reduced by using the inertial suspension. The value of g obtained for Bogota reflects this by having a greater uncertainty than that obtained from the other sites. The best data was obtained at the BIPM and Middletown sites where a few sets taken on particularly quiet nights show a scatter as small as 0.045 mgal.

This drop-to-drop scatter is, however, not the only random variation in the measured value of g. There are both long term and short term variations in the wavelength of the laser light which are either the result of slow drifts in the servo-locking system or of repeatability of the frequency to which the laser is locked. This can be due to different operating characteristics depending on temperature, locking procedure, etc. Fluctuations of up to ± 2 parts in 10^8 would not be surprising. Another possible source contributing fluctuations is small, long period accelerations of the inertial suspension system due to thermal drifts. At sites where the seismic background

[1] The unit of acceleration used throughout this paper is the milligal (mgal) which is equal to 10^{-5} m/s^2.

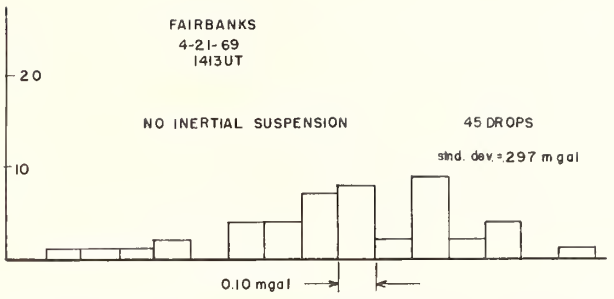


FIGURE 5. *Histogram of data without suspension of reference corner-cube.*

is low enough to take data without the inertial suspension in use a lower scatter among the averages of the data sets is usually observed. These results seem to indicate that the inertial suspension system might contribute a scatter between sets of about ± 1 part in 10^8.

The measured value and its uncertainty are then corrected for several natural and instrumental systematic effects. The vertical gradient of the earth's gravitational field measured at each of the sites (about 0.30 mgal/meter) is used to reduce each stated value to the floor at each particular site. In practice we make this correction in two parts: first we correct the measured value to the zero-velocity position of the dropped object. This correction is given by:

$$g_0 = g_m(1 - \tfrac{1}{12}\gamma T_2^2(R^2+R+1) + \tfrac{1}{3}\gamma T_2 \Delta t(R+1) + \tfrac{1}{2}\gamma(\Delta t)^2)$$

where T_2 is the longer time interval and R is given by $R = T_1/T_2$ where T_1 is the short time interval; γ is the gravitational gradient, g_m is the measured value of gravity and Δt is the initial time interval from release till when the measurement begins. Having done this, a simple gravimetric correction is then made to transfer this value to the floor. At each site the gradient was measured with a relative gravimeter. The uncertainty in this correction is essentially the reading error of the gravimeters and amounted to ± 0.01 mgal for this vertical transfer.

Another correction needs to be applied as a consequence of the finitude of the velocity of light. As the interferometer's arm length changes, the round trip transit time also changes and this requires that a small correction be made which depends upon the ratio of the velocity of the object to the velocity of light. The expression for this correction is

$$\frac{\Delta g}{g} = -\left(\frac{4}{3}\frac{g(T_1+T_2)}{C} + \frac{2V_0}{C}\right).$$

This may also be called a "doppler effect" correction. Since this correction amounts to only about 3 parts in 10^8 both g and V_0 can be calculated well enough from the uncorrected results.

The effects of random errors in the determination of the time intervals are, of course, included in the

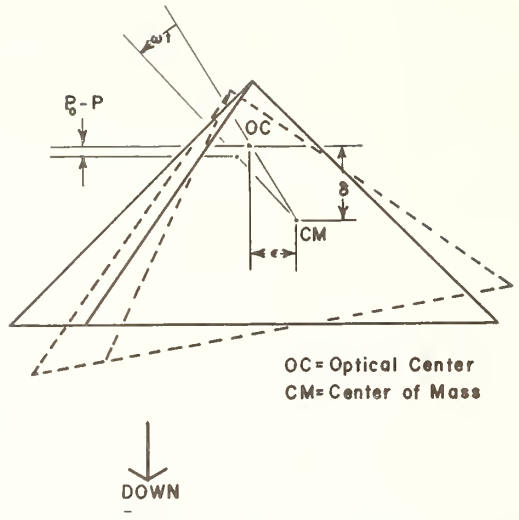


FIGURE 6. *Separation of center of mass and optical center in a corner-cube.*

random drop-to-drop scatter and thus need not be applied as a systematic correction. There is, however, a chance for systematic errors in the time intervals to occur. Careful calibration of the time intervals has been done with adjustments made in cable lengths to assure correct correspondence between the time intervals set on the time delay generator and the actual "on time" of the scaler. The uncertainty in determining the errors of the time of the end points which results from errors in determining the relative position of two 5 ns FWHM pulses on an oscilloscope is estimated to be ± 0.5 ns. Thus while the size of the possible systematic effect due to time errors is estimated to be zero, the uncertainty in that result is estimated to be one part in 10^8 or ± 0.01 mgal.

Errors which would be caused by the possible rotation of the dropped object around a point other than its optical center are minimized by designing it to have its center of mass at the optical center of corner cube.

It has been shown by Peck [5] that rotation about the optical center of a solid corner cube changes the optical path by an amount whose lowest order angular dependence is in the fourth power of the angle of rotation. For small rotations about points other than the exact optical center the change in the optical path is the same as the geometrical effect of translating the corner cube the appropriate distance. Thus if we consider the effect of rotating about a point separated from the optical center as shown in figure 6 the change in optical path for a small rotation given by α is:

$$d = \epsilon\alpha + (\delta\alpha^2/2)$$

where ϵ and δ are the horizontal and vertical components of the separation of the optical center and the point of rotation (center of mass). If the angular velocity is ω then this displacement as a function of time, t is

$$d = \epsilon\omega t + (\delta\omega^2 t^2/2).$$

The false acceleration given by $\delta\omega^2$ is minimized by controlling δ and ω. The term proportional to ϵ acts like an "initial velocity" and as such does not effect the determination of g. This term would, however, be important were the interferometer used to measure either velocity or displacement. The dropped object in this instrument is adjusted to have its center of mass at the optical center of the corner cube. The displacement, δ, is zero to within ± 0.05 cm. The angular rotation of the dropped object can be observed by looking at a reflection from a mirror polished on its top side using the telescope shown in figure 2. This rotational velocity is always less than 5×10^{-3} radian/second. The resulting maximum value of this false acceleration is about one part in 10^9 of g and is therefore negligible.

The effects of errors in either establishing the vertical or collimating the light beam are to shorten the distance corresponding to one fringe and thus to make the measured value of g higher than the true value. These effects are given by approximately,

$$\delta g/g \simeq \alpha^2/2.$$

The angles are checked carefully and known to be less than 10^{-4} radians but since the effects can only be of one sign we have included a correction of

$$\delta g = -0.005 \pm 0.005 \text{ mgal.}$$

The effects of nongravitational forces such as those due to electric and magnetic fields need also to be considered. Our experiment is designed to minimize all electrostatic forces on the dropped object. Static charges on either the dropped object or the inside of the vacuum chamber are neutralized by contact with ground. Contact potential differences are minimized by making the contacting parts from the same type of material. Furthermore, the front face of the corner cube is coated with a conductive antireflection coating and the glass tube of the vacuum chamber is coated on the inside with chromium to eliminate static charging. The indicated effects are small enough that the possible error will be included in the estimate for magnetic effects.

Magnetic forces acting on the dropped object itself are similarly minimized by the design of the instrument. Thus, while there is a piece of soft iron incorporated in the dropped object whose remanant field could interact with local vertical gradients, the effective dipole strength of the soft iron and these gradients themselves are so small that this effect is certainly less than one part in 10^8 of g. The strength and polarity of the iron piece depends on the magnetic field of the holding solenoid as well as the local value of the earths field. It has been found that changing the direction of the field of the holding solenoid does not change the sign of the effective remanant field but changes its magnitude by a factor of two. Since tests of the instrument without the seismometer indicate no difference between data

taken with the solenoid having the two polarities it is certain that any effect is less than ± 0.01 mgal.

Another possible magnetic effect is eddy current damping of the dropped object caused by the piece of magnetized soft iron falling through the conducting cylinder. In Faller's Princeton experiment this effect contributed a correction of $+1.0$ mgal to his final value. In this experiment several important design differences were used so as to make this effect negligible. Merely the fact that the enclosing conducting cylinder is chromium with a thickness of only 10^{-4} of that in Faller's experiment is enough to make this effect negligible in our case. Coupling this with the fact that the diameter is several times larger, even the fact that the velocity is considerably higher does not make the effect large enough to worry about. Considering all these things we have assigned a ± 0.01 uncertainty to a zero value correction for all electrostatic and magnetic effects involving the dropped object directly.

Air resistance will affect the measured value of g though we operate under such high vacuum that this correction is very small. Several estimates of the magnitude of effects due to air resistance as well as extrapolation of measured effects from low vacuum to the high vacuum at which we operate indicates that we should make a correction given by:

$$\Delta g = (+0.001 \pm 0.0005) \times 10^7 \times P \text{ mgal}$$

where P is the pressure in torr. Our operation is usually at pressures of 5×10^{-7} torr so that this effect is usually less than $+0.005 \pm 0.0025$ mgal.

The largest source of uncertainty is due to an effect caused by the magnetic field of the solenoid acting on magnetic materials in the seismometer to which the reference corner cube is attached. This caused the measured value of data for the two directions of the solenoid to differ by several parts in 10^8 at most sites. The effect is due to the interaction of the solenoid field with the seismometer's large Invar spring. When the solenoid is turned off the inertial system accelerates back to its original position and the spurious acceleration of the reference corner cube is included in the measurement. The Invar spring has a large permanent magnetization so that it is reasonable to assume the effect can be considered that of a non-uniform field acting on a long magnetic dipole. Thus we assume the effect reverses sign for the two directions of the field and that the actual value of g should be taken to be the average of the values obtained for the two field directions. Measurements with the seismometer pinned confirm this interpretation, however to account for a possible systematic error in this account we feel that an uncertainty of ± 0.03 mgal must be assigned until the effect is better understood. This is called the "Seismometer Magnetic" effect in table 1.

Table 1 is exemplary of the summary of corrections and estimates of uncertainties for the various possible systematic effects, both natural and instrumental. Differences in the value of the "Net Correction" from site to site are due mostly to differences in the vertical gradient with the NBS value affected by a difference in wavelength for the laser, also.

Table 1. *Typical Corrections and Systematic Effects and Their Estimated Errors*[a]

Source	Correction (mgal)	Est. Uncertainty (mgal)
1. Laser Wavelength	-0.094	± 0.020
2. Direction and Collimation	-0.005	± 0.005
3. Time Intervals	0.00	± 0.010
4. Gravitational Gradient	$+0.432$	± 0.010
5. Velocity of Light	-0.028	± 0.001
6. Air Resistance	$+0.010$	± 0.005
7. Electrostatic and Magnetic	0.00	± 0.010
8. Seismometer Magnetic	0.00	± 0.030
Net Correction	$+0.315$	± 0.041

[a] These data are for the BIPM site, Sèvres A.

The correction to the laser wavelength comes about because all computations of g use a value of λ which was not correct for the laser we used. The actual wavelength was measured by Rowley and Wilson of the NPL and the correction shown is the difference between values of g using the two different values of λ. The uncertainty of the correction comes from the uncertainty in the measured value of λ.

Table 2 is a summary of all final results with a description of the site which is sufficient to locate the site uniquely. The values are given at the floor in the rooms indicated.

Table 3 gives other important information about the data for each site including the average measured value, the number of sets of data, the standard error of the measured value and the value of the net correction and its uncertainty. The standard error of the measured value is obtained by quadratically combining the standard errors for the data taken with the two different directions of the solenoid field. The resulting final value and its uncertainty are also given for each site. The NBS site was first visited and because of difficulty with the laser operation we were unable to take more than five sets of data. In addition, the operating wavelength at that time was not so precisely known.

Table 4 is a summary of the most recent absolute measurements of the acceleration of gravity made at sites occupied by our instrument. For the purposes of comparison the results of Faller and Tate together with that of the present experiment at NBS have been transferred to the national gravity base in the Department of Commerce Building (east elevated pier in the gravity room). A gravity difference between NBS-2 and the Commerce Base (CB) of -2.97 ± 0.03 mgal as given by Tate is used in transferring both Tate's value and that of Hammond-Faller experiment. The transfer of Faller's Princeton measurement was made by applying this transfer value plus a separate transfer between Princeton and NBS-2.

The Potsdam correction given is based on a value

TABLE 2. *Summary of measurements with the Hammond-Faller apparatus*

Site	Location	g at floor (mgal)
National Bureau of Standards Gaithersburg, Md.	Rm. 01, Bldg. 202 (NBS-3)	980 102.394±0.055
National Physical Laboratory Teddington, England	Rm. B-17, Bushy House (BH)	981 181.930±0.042
Bureau International Des Poids et Mesures Sèvres, France	Salle 1, BIPM Laboratory (Sevres A)	980 925.960±0.041
Air Force Cambridge Research Laboratories Bedford, Mass.	Pier 1, Seismic Facility	980 378.671±0.042
Institute of Geophysics, University of Alaska Fairbanks, Alaska	Rm. #1, Patty Bldg.	982 234.953±0.042
Universidad Nationale de Colombia Bogota, Colombia	Quarto 111, Edificio Matematica y Fisica	977 390.015±0.087
University of Denver Denver, Colo.	Rm. 8, Science Hall	979 597.708±0.042
Scott Laboratory of Physics Wesleyan University, Middletown, Conn.	Rm. 7B, Scott Lab. (Middletown A)	980 305.306±0.041

TABLE 3. *Summary of data obtained with this instrument (all values in mgal)*

Site	Number of sets*	Average value	Stnd. error	Net corr.	Est± error	Final value
NBS-3	5	980 102.204	±0.015	+0.190	±0.053	980 102.394±0.055
NPL	19	981 181.602	±0.009	+0.328	±0.041	981 181.930±0.042
Sèvres A	27	980 925.645	±0.008	+0.315	±0.041	980 925.960±0.041
AFCRL	22	980 378.355	±0.011	+0.316	±0.040	980 378.671±0.042
Fairbanks	16	982 234.682	±0.010	+0.271	±0.040	982 234.953±0.042
Bogota	38	977 389.745	±0.071	+0.270	±0.041	977 290.015±0.087
Denver	30	979 597.410	±0.012	+0.298	±0.041	979 597.708±0.042
Middletown	100	980 304.994	±0.007	+0.312	±0.041	980 305.306±0.041

* Each data set consists of 50 drops ($\frac{1}{2}$ hours data). A single tidal correction is applied to each set of data.

TABLE 4. *Summary of recent measurements of the acceleration due to gravity*

Location and Author	Value implied at site indicated (mgal)	Difference from Potsdam system (mgal)
NBS-3; Hammond–Faller	CB: 980 104.234±0.07	−13.77 ±0.08
NBS-2; Tate[a]	CB: 980 104.77 ±0.45	−13.23 ±0.45
Princeton, N.J.; Faller[b]	CB: 980 103.8 ±0.7	−14.2 ±0.7
NPL; Hammond–Faller	BFS: 981 181.86 ±0.05	−13.60 ±0.80
NPL; Cook (Revised 1969[c])	BFS: 981 181.81 ±0.31	−13.66 ±0.14
BIPM; Hammond–Faller	SA: 980 925.960±0.041	−13.800±0.08
BIPM; Sakuma[d]	SA: 980 925.949±0.006	−13.811±0.070

[a] Reference 6
[b] Reference 7
[c] Reference 8
[d] This is Sakuma's most recent result first presented at this Conference.

of g at Potsdam of 981 274 mgal. The Comité International des Poids at Mesures at its October 1968 session has recommended that as of 1 January 1969, g (Potsdam) be reduced to 981 260 mgal, thus defining a revised Potsdam system which is in better agreement with modern absolute measurements [9].

The results clearly indicate some of the inconsistencies in the presently accepted transfer values from site to site. In particular the difference between the BFS and Sèvres A sites indicated by our measurements is:

$$g(\text{BFS}) - g(\text{Sèvres A}) = 255.91 \text{ mgal}$$

while that given by Cook [10] as the result of a number of transfers with relative gravity meters is:

$$g(\text{BFS}) - g(\text{Sèvres A}) = 255.71 \text{ mgal}.$$

The stated uncertainty of this latter value is 0.06 mgal so the difference of 0.20 mgal is quite significant. Because our value is the result of two absolute measurements without the drift or calibration problems of relative instruments we feel the uncertainty of such a transfer value is ± 0.02 mgal.

It seems unreasonable to persist very much longer in maintaining a World Gravity Network based on a measurement (or a correction to that measurement) made more than 60 years ago at a site which is not always easy for scientists in much of the Western world to visit. If, in fact, a single site must be considered as the basis of a World-wide Network then that site should probably be the one at the BIPM at which two corroborative absolute determinations have been made. These two measurements have been made by independent experimenters using methods which have fundamental differences and the agreement of the values obtained reinforces the confidence with which the results may be accepted.

It is, of course, no longer necessary to have a network based on the value at only one particular site. Instead, any site at which an absolute measurement has been made either with the Hammond-Faller instrument or any other either in the (recent) past or in the future can be considered a base-station with a "quality" indicated by the precision of the absolute measurements made at the site. Thus the BIPM site would probably be the best site among a large number of calibrated sites throughout the world.

The result, then, of the work presented here has been threefold: (1) An instrument has been built which has produced new absolute values of the acceleration of gravity at a number of international sites of importance from a both geodetic and a standards point of view. (2) For the first time several absolute measurements have been compared with those of one portable instrument offering useful corroboration of all of the involved measurements. (3) The success of the instrument and therefore of the method used leads the way for the development of other instruments of similar nature. Instruments could be built with less emphasis on portability having the required higher precision for the investigation of possible geophysical and cosmological effects and without much loss in precision, portable devices of smaller dimensions could be built for use in extensive geodetic work. In terms of simplicity, adaptability, and proven capability, the laser interferometer method of measuring the acceleration of gravity has much to recommend it as the basis of future measurements of this important physical quantity.

References

[1] Hammond, J. A., and Faller, J. E., Laser-interferometer system for the determination of the acceleration of gravity, IEEE J. Quantum Electronics QE-**3**, No. 11, 597–602 (1967).

[2] Faller, J. E., The precision measurement of the acceleration of gravity, Science **158**, 60 (1967).

[3] Hammond, James A., A Laser-Interferometer System for the Absolute Determination of the Acceleration of Gravity, JILA Report #103, Joint Institute for Laboratory Astrophysics, Boulder, Colorado.

[4] Longman, I. M., Formulas for computing the tidal accelerations due to the moon and the sun, J. Geophys. Res. **64**, 2351 (1959).

[5] Peck, E. R., Theory of the corner-cube interferometer, J. Opt. Soc., **38**, No. 12, 1015–1024 (Dec. 1948).

[6] Tate, D. R., Acceleration Due to Gravity at the National Bureau of Standards, J. Res. Nat. Bur. Stand. (U.S.), **72C** (Eng. & Instr.) No. 1, 1–20, (Jan.–Mar. 1968). (Available as NBS Monograph 107 from the U.S. Govt. Printing Office, Washington, D.C.)

[7] Faller, J. E., An Absolute Interferometric Determination of the Acceleration of Gravity, Thesis, Princeton University, 1963.

[8] Cook, A. H. and Hammond, J. A., The Acceleration Due to Gravity at the National Physical Laboratory (Note), Metrologia **5**, No. 4, 141–142 (Oct. 1969).

[9] Terrien, J., News from the Bureau International des Poids et Mesures, Metrologia **5**, 68 (1969).

[10] Cook, A. H., A New Absolute Determination of the Acceleration Due to Gravity at the National Physical Laboratory, England, Phil. Trans. Roy. Soc. London, Ser. A, **261**, 211–252 (1967).

A Determination of the Acceleration Due to Gravity at Kakioka, Japan

Tadao Inouye, Tomiji Kitsunezaki, Osamu Senda, and Koichi Ando

National Research Laboratory of Metrology, Itabashi, Tokyo, Japan

This measurement is based on the free-fall method using a 1 metre graduated scale. The obtained value of gravity is 9.79964_8 m/s^2 with the assigned error of 1.0×10^{-5} m/s^2. This value is 15.3×10^{-5} m/s^2 smaller than the value based on the Potsdam system.

From analysis of the observational data, it was found that some of the experiments were affected by microseisms.

Key words: Acceleration due to gravity; free-fall method; Potsdam system.

1. Introduction

An absolute measurement of the acceleration due to gravity in Japan has been carried out at the gravity station of the NRLM at Kakioka, about 70 km to the north-east from Tokyo. According to the classification [1] by Cook, this method is called the multiple position free-fall method.

2. Equipment for the Measurement [2]

The equipment for the measurement is shown in figure 1, and its scheme, in figure 2.

Falling scales used in the measurement, one made of invar and another of fuzed quartz, are about 1 m in length and the graduations are ruled every 1 cm. Their lengths have been calibrated interferometrically.

Either of these scales falls freely in a brass cylinder of 3.5 m in height and 0.3 m in inner diameter and evacuated to 1×10^{-5} torr.

Before falling, the scale is hung with a nylon string at the top of the cylinder, the graduated surface is kept vertical and faced to north, its normal line is made to coincide with the optical observing axis, and the original graduation line at the bottom of the scale is located in the observing field.

At the beginning of a fall, the nylon string is cutoff by a small electric heater. A special timing device is arranged to control the electric power supply to the heater and the operation of a running roll film camera.

At the end of a fall, the scale is retarded and stopped by a braking mechanism with a braking path of 1 m long. The maximum retarding acceleration of the scale is lower less than five times the gravitation.

During the experiment, the operation of the rotary pump and main other sources of vibration are stopped. The evacuating system is so designed that the oil diffusion pump has enough back volume to maintain a specified vacuum without use of a rotary pump during observation.

3. Reading of the Fall Time [3]

The graduated surface of the scale is illuminated by a horizontal light beam. In the course of falling motion of the scale, the graduations enter the observing field successively. When a graduation passes through the illuminating light beam the reflected light flux decreases, and a photomultiplier, located behind the image of the graduation, generates an electric signal each time when a graduation runs across the observing field.

The electric signal is reproduced on the Y axis of a cathode-ray-tube (CRT) display through a delay line, while a trigger pulse corresponding to each electric signal is generated by a pulse generator and transmitted to the trigger terminal of the CRT display. Thus profiles of the signals are traced on the CRT screen and recorded in the running roll film camera. The brightness of the profile of the signal on the CRT screen is modulated by 0.2 μs time mark, with which the time intervals between the peaks of successive recorded profiles can be read with a time division of 0.1 μs. The display and time marker system is illustrated in the figure 3. Figure 4 shows a part of the records. There are beginnings of bright dots at 10 μs intervals followed by a certain number of bright dots representing the figure in the 10 μs decimal.

4. Data of the Gravity Experiments [4]

The relation between the distance y_i and the time t_i of fall corresponding to the ith graduation is expressed by the equation

$$y_i = A + Bt_i + \tfrac{1}{2}gt_i^2 \qquad (1)$$

where A and B are the constants depending on the

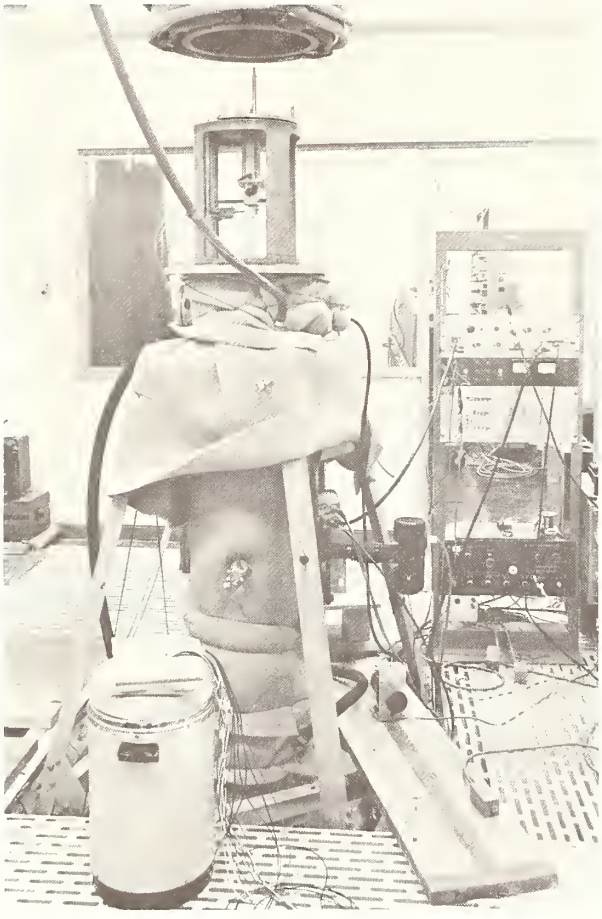


FIGURE 1. *Equipment for the measurement of gravity.*

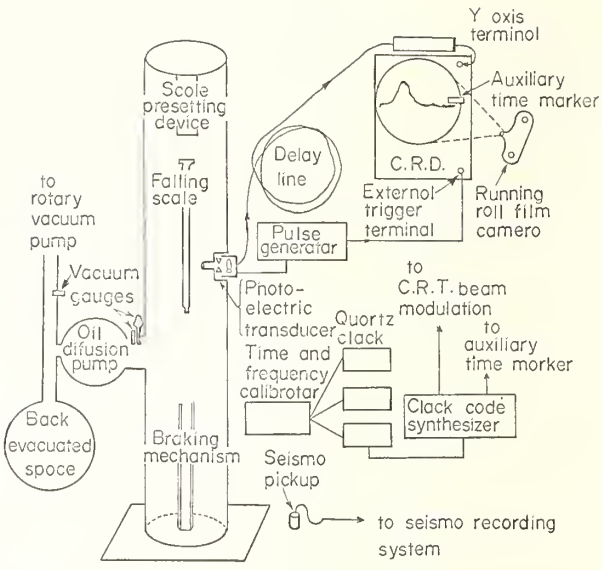


FIGURE 2. *Schematic diagram of g-measurement system of the NRLM.*

origin of the coordinates. If the observation is made under a uniform gravitational field and there are no other effects which affect the falling scale, "g" will coincide with the acceleration due to gravity. Corresponding to each fall, g can be determined by applying the method of least squares.

In the treatment of obtained data, the weights for y_i and t_i are assumed to be equal since their precisions are estimated to be nearly equal, that is 0.15 μm and 0.14 μs.

Twenty-three falls with the invar scale and 15 falls with the quartz scale have been performed during 1965 to 1967. The obtained values of g, together with their respective 95 percent confidence limits, are presented in table 1 and figure 5. The thermal expansion coefficient and the compressibility of the scales are of course taken into account.

5. Exclusion of Inconsistent Data

By applying a regression analysis with respect to t_i, it is found that some of the data are significant to 1 percent for the existence of the third order component of t_i. These data are marked with **.

As a way to find the cause, the symmetry of the recorded signal profiles is examined. Those data in the table 1 for which the peak of the profile is shifted by more than 5 percent of the half-value width from its center are marked with ^. In these data it is not improbable that t_i has not been estimated with suffieient exactitude, and that the existence of the third order component of t_i becomes significant. The four experiments in table 1 (a), where the marks ** and ^ are overlapping, have therefore been excluded as a matter of inconsistency.

Asymmetric profiles will probably be due to deviation of the axis of the scale from the vertical direction caused by a microseismatic force or the interaction between the terrestrial magnetism and the residual magnetism of the invar scale. According to seismographs during experiments, the peak-to-peak amplitude of seisms amount to as much as 0.4 μm in some cases.

6. Results

After exclusion of the four data as discussed in the preceding paragraph, the two mean values corresponding to the invar scale and the fused quartz scale are 9.79964_4 m/s^2 and 9.79964_7 m/s^2 respectively. Since there is no significant difference between these two mean values, it can be concluded that there exists no systematic error due to difference in scales, and the value 9.79964_6 m/s^2 is obtained as the overall mean with the 95 percent confidence limit of 0.7×10^{-5} m/s^2.

The present experiment has been carried out at a point about 88 cm higher than the floor level of the gravity station[1] where the value in the Potsdam

[1] The NRLM gravity station in the yard of Kakioka Geophysical Observatory of Tokyo University, Ibaraki prefecture, Japan (Latitude: 36°13.'7 N, Longitude: 140°11.'7 E, 32m above the sea-level).

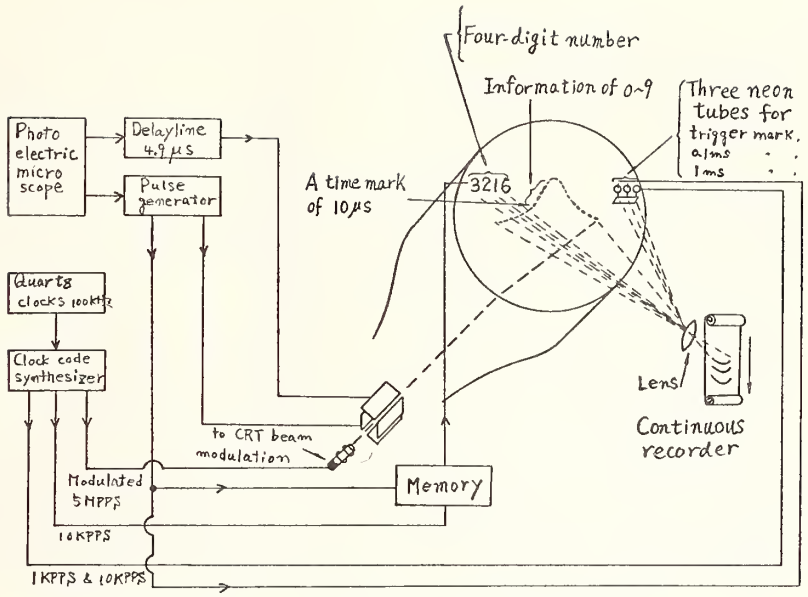


FIGURE 3. *Schematic diagram of the falling time recorder.*

system had been determined in 1963. After correction for this height difference, the reduced value is 9.79964_8 m/s^2, which is smaller than the value in the Potsdam system by $15._3 \times 10^{-5}$ m/s^2.

6.1. Sources of Bias

As for possible sources of systematic error in the measurements, there are four items worth being examined: (1) the frequency characteristics of delay time in the delay line, (2) the flatness of the graduated surface, (3) the longitudinal vibration of the scales and the evacuated chamber by starting the fall motion of scales and (4) the vertical braking action by electromagnetic force between the cylindrical brass chamber and the invar scale.

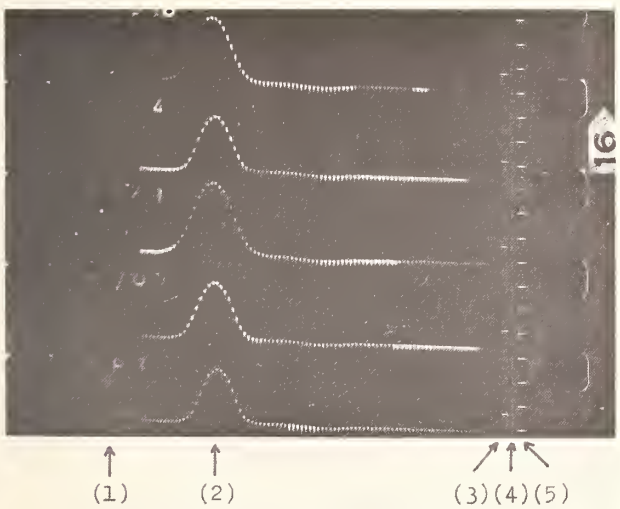


FIGURE 4. *Part of the experimental records.*
(1) Four-digit numbers indicate the falling time of the graduations in units of 0.1 ms.
(2) The signal profiles transduced photoelectrically from the graduations. On the trace, 0.2 μs period dots are produced by intensity modulation. The left end of a train of bright dots is 10 μs time mark, and the length of the train represents the figure in the 10 μs decimal.
(3) Trigger time mark (bright mark).
(4) 0.1 ms time mark.
(5) 1 ms time mark.

Variation of the delay time corresponding to frequency band from zero to 1 MHz, required to transmit the electric signals of the graduations, has been found to be smaller than 0.05 μs and negligible.

Variations in the flatness of the graduated surfaces of the two scales has been found to be much smaller than the depth of the field of the observing lens, so that the bias in the measurements is smaller than 0.03 μm, even if the incident angle of illuminating light may slightly differ from zero.

As for the longitudinal vibration of the scale and the cylinder, the amplitude larger than 0.1 μm have not been observed. The use of the small electric heater in cutting the nylon string gives rise to initial shock on neither the scale nor the cylinder.

The error due to the vertical electromagnetic force may reduce the g value by less than 0.6×10^{-5} m/s^2 according to an estimate based on the measured value of the residual magnetism along the axis of the invar scale and an assumption of the interlinkage area of the magnetic flux with the brass cylinder.

6.2. Representation of the Accuracy

By reason of a difficulty in exact evaluation of the error due to the electromagnetic force, no correction has been applied to the data in the table 1 (a). In place of the possible correction, the value of the statistical confidence limit assigned to our final result, $0._7 \times 10^{-5}$ m/s^2, should be broadened by one half the estimated systematic error and finally be evaluated as $1._0 \times 10^{-5}$ m/s^2.

(a) Invar scale

Experiment Number	Measured Value (m/s^2)			95% Confidence Limit (ppm)
93	9.79965_6			$1._1$
94	9.79966_3			$1._1$
95	9.79965_4			$0._9$
96	9.79965_0			$1._0$
97	9.79961_7			$1._0$
98	9.79968_2			$1._1$
100	9.79963_0			$1._4$
101	9.79964_5			$1._3$
103	9.79964_8			$1._0$
104	9.79964_0			$1._1$
105	9.79962_4		^	$1._0$
106	9.79964_9		^	$1._1$
107	9.79967_5		^	$0._9$
108	9.79966_4		^	$1._1$
112	9.79963_7	**	^	$1._2$
113	9.79963_1		^	$0._9$
114	9.79961_9		^	$0._9$
116	9.79957_5	**	^	$1._1$
117	9.79963_9			$1._0$
118	9.79960_3	**	^	$0._8$
119	9.79961_6	**	^	$0._9$
120	9.79962_3			$0._7$
121	9.79963_3			$0._6$

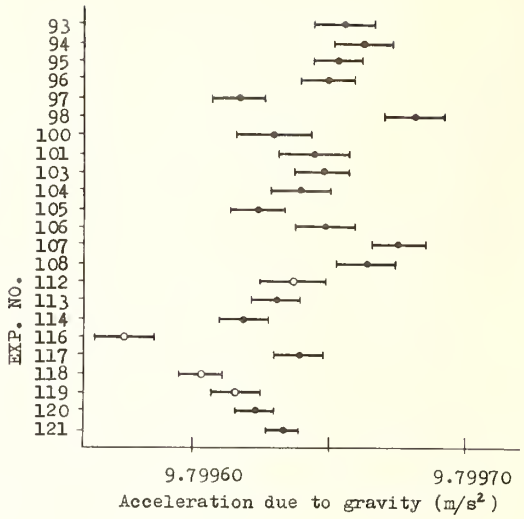
(b) Fused quartz

9	9.79963_6		^	$0._8$
10	9.79967_0			$0._7$
12	9.79963_3	**		$0._6$
13	9.79965_8	**		$0._6$
14	9.79969_5		^	$0._8$
16	9.79965_0			$0._8$
17	9.79959_4		^	$1._0$
18	9.79966_0			$0._9$
19	9.79964_4			$0._7$
20	9.79962_1			$0._7$
21	9.79964_4			$0._7$
22	9.79966_0			$0._7$
23	9.79965_3			$0._7$
26	9.79963_8			$0._7$
27	9.79964_6	**		$0._5$

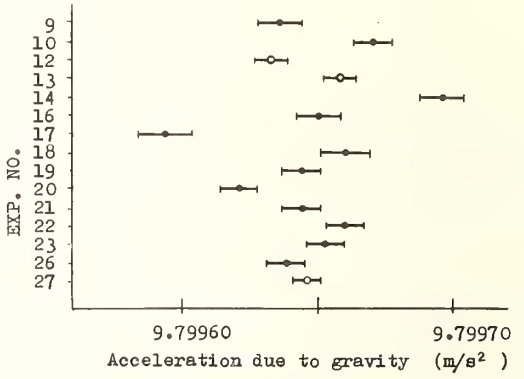
** : 1% significant data

^ : data with asymmetric profiles

TABLE 1. *Data of the gravity experiments*



(a) Invar scale



(b) Fused quartz scale

FIGURE 5. *Measured values: (a) invar scale, (b) fused quartz scale.*
⊢●⊣: 95 percent confidence limit
○: significant to 1 percent for the 3d order component.

7. Acknowledgments

The original idea of this measurement and much of the initial works are due to Dr. A. Sakuma of the International Bureau of Weight and Measures. The authors wish to express their thanks to Professor R. Yoneda of Kogakuin University, Dr. K. Yamamoto, Director of the NRLM, Mr. M. Tamano and Dr. Y. Tomonaga, Former Directors, for their valuable suggestions and encouragements.

8. References

[1] Cook, A. H., Metrologia **1**, 84 (1965).
[2] Sakuma, A., KEISOKU (Journal of the Society of Instrument Technology, Japan) **9**, 511 (1959).
[3] Kitsunezaki, T., NIHON TOKEI GAKUKAI SHI (Journal of the Horological Institute of Japan) **37**, 5 (1966).
[4] Inouye, T., Senda, O., Ando, K., and Kitsunezaki, T., OYOBUTSURI **39**, to be published in August, 1970.

The Absolute Determination of the Gravitational Acceleration at Sydney, Australia

D. L. H. Gibbings, J. B. Patterson, and G. A. Bell

CSIRO National Standards Laboratory, Sydney, Australia

An absolute measurement of the gravitational acceleration "g" has been made at the National Standards Laboratory, Chippendale, N.S.W., Australia.

The determination was made by studying the free motion of a body projected vertically upwards in a vacuum and the time between its initial and final passages through two horizontal planes of known vertical separation was measured.

The measured value of g at a point 1.2 metres above the floor in room B.37 of the National Standards Laboratory is

$$9.7967134 \text{ m/s}^2.$$

The corresponding value at floor level at the BMR gravity station is

$$9.796717 \text{ m/s}^2.$$

Key words: Absolute measurement; acceleration due to gravity; gravity.

1. Introduction

An accurate knowledge of gravitational acceleration is essential in any standardizing laboratory responsible for the maintenance of the standards of force and pressure and is also required in the realization of the basic electrical units. Since Australia is situated many thousands of kilometres from the site of the nearest absolute determination of "g" and is thus dependent on ties over very large intervals of space and time, it was considered essential to make an absolute determination of "g" at the National Standards Laboratory, Sydney.

The site chosen for the determination is in room No. B.37, approximately 1.6 metres from the plug marking the station used by the Bureau of Mineral Resources as the site for a number of pendulum and gravity meter observations. The absolute value found can therefore be fed into the world network with relatively little further work and the value compared with those found in other modern absolute determinations.

It was decided that an absolute determination in which the uncertainty in the determined value exceeded 1 part in 10^6 was pointless, and that it would be desirable to achieve an accuracy approaching 1 in 10^7 if possible.

2. Principle of the Method

The advantages and disadvantages of various methods of measuring "g" including reversible-pendulum, free-fall and rise-and-fall experiments have been discussed by Cook [1] and it was decided that a rise-and-fall experiment in which a study is made of the motion of a body projected vertically upward in a vacuum would probably give the most accurate result despite the obvious technical problems involved.

The arrangement of the experiment is shown in figure 1. The beam splitter of a Michelson type interferometer divides the collimated beam of white light from a zirconium arc. A hollow cube-corner reflector is projected vertically up and reflects the vertical beam of light through 180° on to the horizontal reference mirror whence it is reflected back through the cube-corner to the beam splitter and focused on to the detector.

The light which leaves the beam splitter horizontally is turned through 90° and part is reflected at the top surface of a length bar and part at a mirror optically contacted to the lower end of the bar.

When the optical path of the rays reflected from either end of the length standard is equal to the optical path through the corner reflector the recombined rays form an interference pattern in white light which is detected by the photo-diode and displayed on the CRO screen. Thus the distance traveled between signals from the upper and lower stations is determined precisely in terms of the length standard.

2.1. The Length Measurement

The bar used for the most accurate measurements which have so far been made is of fused silica and is of length 0.41 metre. The standard is in the form of a cylindrical tube with flat parallel ends. To eliminate residual errors in the parallelism a pair of fused silica wedges of angle 2 s of arc were made

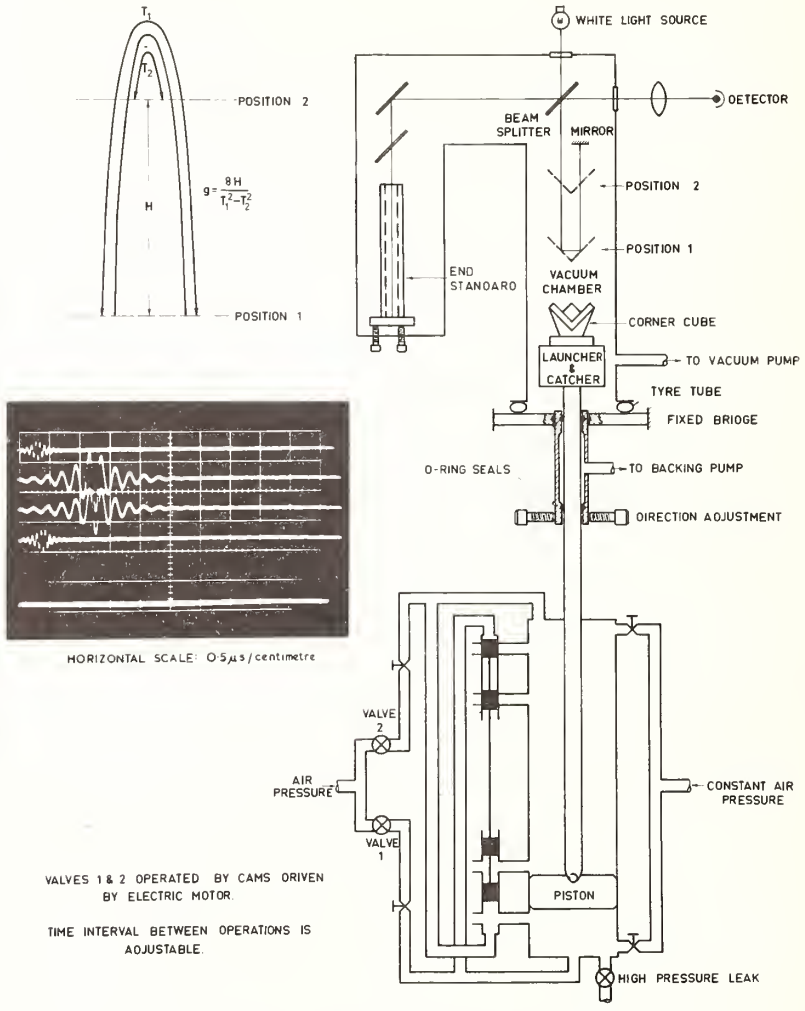


FIGURE 1. *General arrangement (schematic).*

and their orientations relative to the standard adjusted to compensate the residual parallelism error of the standard. This composite wedge was optically contacted to the standard. A fused silica flat with a metallized zone in the centre was optically contacted to the other end of the standard which was measured in vacuo at 20 °C in terms of the ^{86}Kr standard radiation using a Kösters type interferometer. The determination of the length of the standard at 20 °C is estimated to be in error by not more than 1 in 10^7.

2.2. Time Measurements

The intervals between consecutive passages of the corner reflector through each measuring plane must be measured in 10^{-8} s if the required accuracy in the measurement of "g" is to be attained. This is achieved by timing the occurrence of the central fringe in each white light interference pattern. In each measurement four such patterns occur and the intervals between the first and fourth patterns and between the second and third patterns are measured.

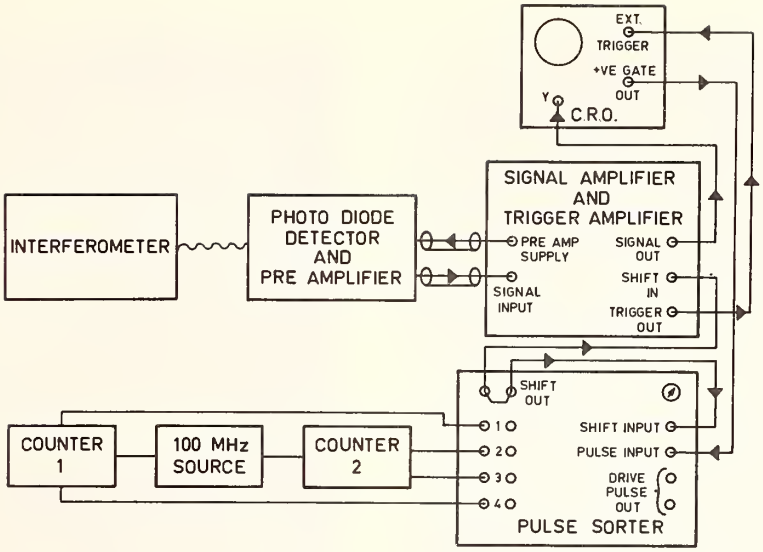


FIGURE 2. *Schematic diagram of timing circuits.*

The general arrangement of the timing circuits is shown in figure 2.

The pulses from the detector are amplified and fed to the external trigger input of a Tektronix 551 oscilloscope. Each pulse triggers the sweep circuit of the oscilloscope which in turn produces a positive gate pulse. This positive gate pulse is utilized to start and stop the counters as required. The time interval between the crossing of a graticule line by the sweep and the front edge of the positive gate pulse is constant within 10^{-8} s. Therefore the constant but unknown delay between the sweep and the positive gate pulse drops out when measuring the separation of the central fringes in the two patterns.

The counters are supplied by a 100 MHz signal from a standard source and are controlled by the positive gate pulse from the CRO.

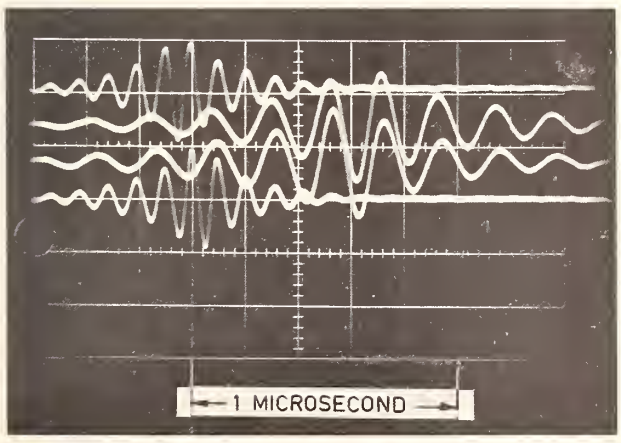


FIGURE 3. *Typical CRO trace of white-light fringe pulses.*

In order that the pulses shall be fed to the counters in the correct sequence they are passed through a pulse sorter and in order to separate the traces vertically a shift voltage is applied 30 μs after each sweep. A typical trace is shown in figure 3.

FIGURE 4. *The corner-reflector resting on a mirror to show launching and catching surfaces.*

3. Components

3.1. The Corner-Reflector

This comprises three fused-silica aluminized mirrors adjustably mounted in an aluminum-alloy body and is shown in figure 4. The three mirrors are flat to 0.025 μm and are adjusted to be mutually perpendicular to 0.2 sec using a Michelson interferometer and a laser source (fig. 5).

It is essential that the centre of mass of the moving body should coincide with the optical centre of the corner reflector. This was achieved using an autocollimator with its axis horizontal and suspending

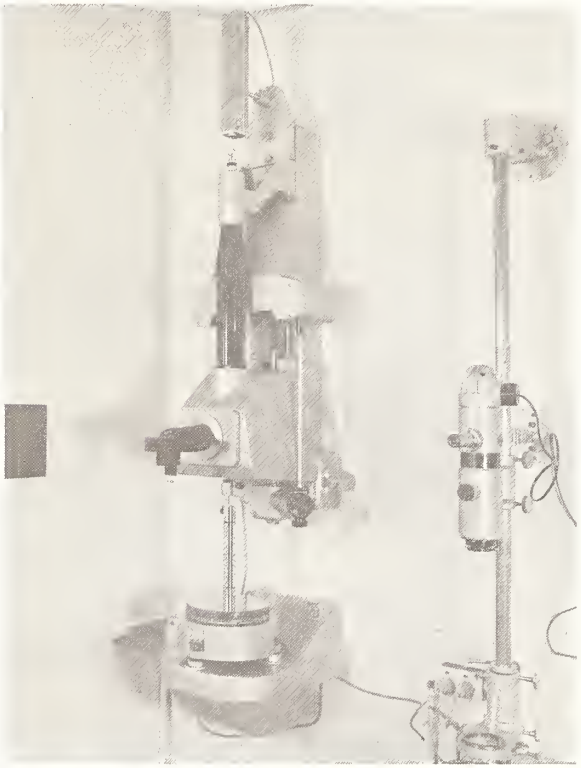


FIGURE 5. *Laser interferometer for adjusting corner-reflector and length standards.*

the corner-reflector from a piece of adhesive tape fixed to each face in turn and adjusting the trimming screws until each suspended face was vertical.

Three vertical rods on the top surface of the catapult support the flat annular surface of the corner-reflector body and the thrust is transmitted by them when the catapult is operated.

The conical lower section of the corner reflector body serves to guide it into its correct position when it descends into the catcher (fig. 6).

3.2. The Catapult

The primary requirement of the catapult is to launch the corner reflector accurately vertically at a controlled velocity and to create a minimum mechanical disturbance.

Mainly on aecount of this last requirement a pneumatic system was used. The catapult is located directly beneath the interferometer in a pit in the floor. It comprises a long cylinder fixed to the bottom of the pit and connected to the pneumatic system as shown. It is surmounted by a piston rod which passes through a long bearing fixed to a heavy bridge immediately below the vacuum chamber. The bearing incorporates a vacuum seal and its direction was made adjustable to enable the direction of motion to be set truly vertical.

The operation of the catapult is as follows. Air at a pressure between 2 and 3 atm is admitted to

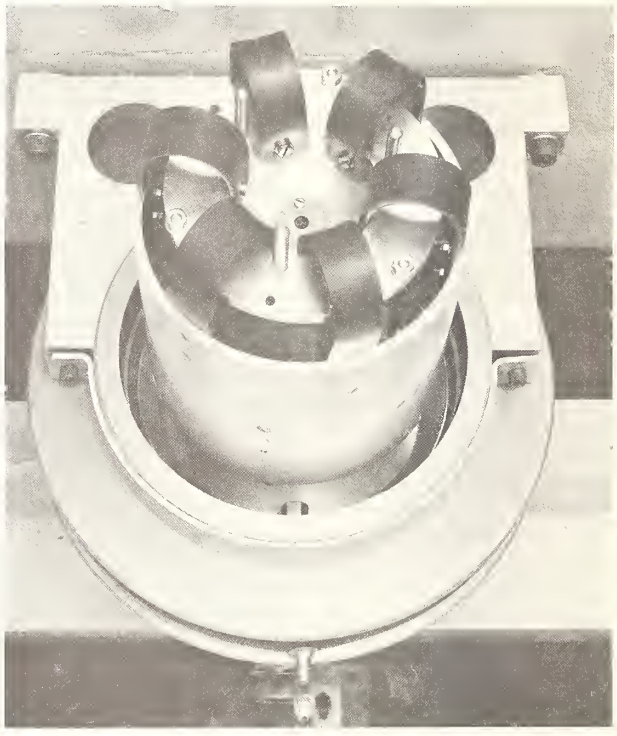


FIGURE 6. *The catcher.*

the lower part of the cylinder and accelerates the piston up the cylinder. When its velocity reaches about 4 m/s the air supply at the lower end is cut off and the upper part of the cylinder is closed and as the piston continues to move up the air pressure above it increases and it is brought smoothly to a halt. At this point the upper part of the cylinder is opened to the atmosphere. The whole of this operation is performed automatically by an electric motor which operates the valves through a series of cams.

To protect the corner reflector from damage at the end of its descent it is essential to have some means of smoothly absorbing its kinetic energy. This is done with the catcher shown in figure 6. It comprises an

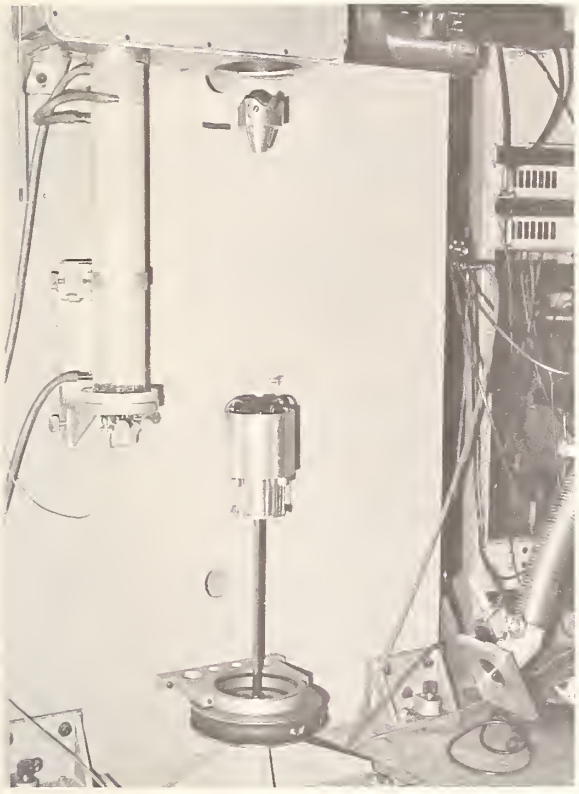


FIGURE 7. *Corner-reflector in flight, vacuum jacket removed.*

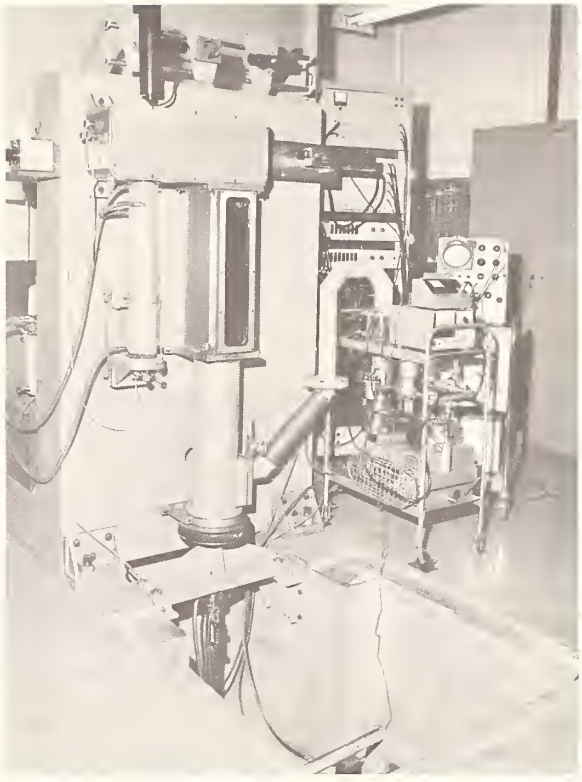


FIGURE 8. *General view of the complete apparatus.*

outer metal sleeve which can slide vertically on the inner member. The sleeve and the inner rings are connected by six neoprene loops. When the piston is retarded and stops at the end of its upward movement the outer sleeve moves up, raising the neoprene straps well above the three supporting points on which the corner reflector rests prior to projection. The descending reflector is caught by the neoprene straps and brought smoothly to rest. The downward motion of the sleeve on the shaft is retarded by a set of dampers and the corner reflector is set down gently on its three-point support in readiness for the next firing. The throwing and catching apparatus, with the corner-reflector in flight are shown in figure 7.

4. Alinement of Interferometer

The correct adjustment of the catapult ensures that the corner reflector moves vertically up and down but it is essential that the reference mirror be accurately horizontal to ensure that the vertical displacement of the body is measured. The mirror was set accurately horizontal using a mercury pool as a horizontal reference and adjusting the mirror to parallelism with it using an autocollimator reading to 0.1 sec.

The length standard is supported on an adjustable mount and alinement can be effected from outside the vacuum chamber. This alinement is very critical as even small departures from correct orientation cause a rapid deterioration in the quality of the fringe pattern.

A general view of the complete apparatus mounted on its concrete support is shown in figure 8.

5. Results

To the middle of June 1970 a large number of throws had been made with the catapult and the whole apparatus adjusted to its optimum condition. In the last week of June, 51 determinations were made over a period of three days which gave the result

$$g = 9.7967134 \text{ m/s}^{-2}$$

with a standard deviation of

$$0.0000019 \text{ m/s}^{-2}.$$

In addition results have been obtained with length standards of different lengths; the effect of varying the pressure in the vacuum chamber has been examined and the effects of an externally applied magnetic field have been studied.

The effect of vibration, particularly that generated by the catapult, has been measured with a Leet-Blomberg vertical seismograph mounted on the concrete block carrying the apparatus. The slight

disturbance due to the catapult decays to close to the general seismic noise level before the corner reflector reaches the first measuring point.

6. Comparison of Absolute Value With "Potsdam" Value

The value assigned to "g" at the BMR station adjacent to the site of the absolute determination is

$$9.796857\ \mathrm{m/s^2}.$$

The value for this station based on the absolute values and the height correction is

$$9.796717\ \mathrm{m/s^2}.$$

The difference 0.000140 $\mathrm{m/s^2}$ is in good agreement with results derived from other recent absolute measurements.

7. Reference

[1] Cook, G. H., Metrologia **1**, 84 (1965).

COMBINED DISCUSSION

E. R. Cohen: I'd like to ask Dr. Hammond whether there has been any thought given to recording the total fringe pattern and then fitting gravity to that, rather than picking on specific stations.

J. A. Hammond: Yes, there has been some thought of that. For instance, videotape recorders are about as high frequency as you can get. And as I recall, when we last looked for those we were hung up because we have a maximum frequency of about 13 or 14 MHz. At that time 4 MHz was about the upper limit of the available recorders. The question that one has to decide first is if you can get much more information doing that. You may not be able to.

A. H. Cook: I would like to emphasize the great importance of Dr. Hammond's and Professor Faller's work, because before they did their experiments no one who did an absolute gravity measurement was able to compare his results with any accuracy with anyone else's, so he had great confidence in his own results but not in anyone else's. That situation has changed. It may not be to the benefit of the experimenters, but it's certainly to the benefit of metrology.

G. W. Series: I just wonder whether there is a correction necessary in any of these experiments to take account of the change of frequency when one is measuring light that is being reflected in a moving mirror. Is there a Doppler effect correction to be brought in here somewhere at this level of precision?

J. A. Hammond: Yes, there is a Doppler effect correction made on my data and Dr. Sakuma's and probably on the other experimenters' too. On my error budget or list of systematic corrections, there is a correction of about 0.03 milligal.

G. W. Series: I'm sorry. I didn't notice that.

An old Cambridge fellow called Newton
Is often supposed an astute 'un.
When he thought of big G,
"It's useless!" cried he.
What a farsighted fellow was Newton.

A. H. Cook

The Experimental Determination of the Constant of Gravitation

A. H. Cook

University of Edinburgh, Edinburgh 9, Scotland

This paper reviews the principal determinations of the constant of gravitation from the work of Cavendish to that of Heyl. The main sources of error are discussed and the reasons for the poor precision of all work so far done are considered. Brief mention is made of the new determination by Beams and his collaborators now in progress at the University of Virginia, and the plans for a new determination at Trieste by A. H. Cook and A. Marussi are described.

Key words: Constant of gravitation; gravitation.

1. Introduction

The constant of gravitation was the first of the so-called fundamental constants of physics to be measured and it is perhaps surprising that although some very careful experimenters have given their attention to the problem since the first experiments of the Hon. Henry Cavendish in 1798 [1], (who developed a scheme of the Rev. H. Mitchell) yet no really significant improvement has been made in the accuracy with which G is known. There are two reasons for this state of affairs—the fact that the force of gravitation is a very small force so that the measurement is a very difficult one; and the irrelevance of a knowledge of G to most practical physics so that there is no technical motive for the measurements. The force of gravity between an electron and a proton is 10^{-39} of the electrical attraction between them, a fact that may emphasise that the gravitational forces that can be achieved in a terrestrial laboratory are very small compared with possible sources of disturbance. Great care and ingenuity are therefore needed to attain significantly higher accuracy than has been achieved in the past, while since the effects of gravitation are purely mechanical, the most precise physical measurements, those of electrical and spectroscopic phenomena, cannot be directly applied as they may be in the determination of all other fundamental constants. Thus it is, as is shown in table 1, that the scatter of determinations of G indicates an uncertainty in our knowledge of G between one and five parts in a thousand as compared with a few parts in a million or better for other fundamental constants.

Although gravitational attraction controls the motions of celestial bodies, natural and artificial, and determines the acceleration to which bodies at the surface of the Earth are subject, yet the actual value of G is not in general needed for any practical purpose. The period, T, of a celestial body in orbit around another obeys Kepler's law

$$T^2 = 4\pi a^3/GM;$$

Table 1. *Determination of* G

Author	Method	G
		$10^{-11} Nm^2/kg^2$
H. Cavendish, 1798	Torsion balance, deflection	6.754
J. H. Poynting, 1891	Common balance	6.698
C. V. Boys, 1895	Torsion balance, deflection	6.658
K. Braun, 1895	Torsion balance, deflection and period	6.658
P. R. Heyl, 1930	Torsion balance period	
	gold	6.678
	platinum	6.664
	glass	6.674
J. Zahradníček, 1933	Torsion balance, resonance	6.659
P. R. Heyl and	Torsion balance [a]	6.676
P. Chrzanowski, 1942	period [b]	6.668
Rose and others 1969	Acceleration	6.674

[a] Hard drawn tungsten suspension.
[b] Annealed tungsten suspension.

T can be measured, as can a, the semimajor axis of the orbit, so that the product GM can be found but not G and M separately. A relative scale of masses may be found from the study of orbits; the masses cannot be related to the terrestrial kilogramme unit; since no experimental relation can be established, the ignorance of G has no practical effect upon the study of the dynamics of celestial bodies.

Only when a mass should be known in terms of the kilogramme is a knowledge of G necessary. Densities of matter in the interiors of the Earth, or planets or stars are found from the gravitational fields and a value of G is required in order to compare them with equations of state for materials determined in the laboratory. The ratio (density/pressure), which has the dimensions (1/velocity^2), is independent of G or standards of mass, but the absolute value of the density reduced to zero pressure does involve the standard of mass or the constant of gravitation. The accuracies with which equations of state and absolute densities can be measured in the laboratory or can be derived from geophysical data now approach one part in a thousand, and the uncertainty of G is beginning to effect comparisons.

The deepest reason for wishing to improve the accuracy of the measurement of G is the relation of the constant of gravitation to the fundamental structure of nature. We know from the work, first of Eötvös [2], then of Dicke [3], that gravitational forces are to a very high degree independent of chemical composition; indeed this result may be used to criticise determinations of G. It is however, implied by certain views of the structure of nature, that the local value of G is controlled by the large scale geometry of the universe; G may therefore possibly vary with time and it may vary with orientation relative to, for example, the disc of the galaxy.

A precision of 1 part in 10^8 is commonly thought to be necessary to detect such effects; that is quite beyond the scope of present experimental methods, but one of the main reasons for experimental work on G seems to me to be to gain experience of the problems involved in exceedingly delicate mechanical measurements so that it may become possible to investigate experimentally the constancy of G in time and direction. I fully expect that no variation will be found, especially since Sakuma [4] has found no significant change in the absolute value of gravity at Sèvres, but the question is one of experiment and is not to be answered hypothetically.

2. Methods for the Determination of G

2.1. Cavendish and His Successors

The torsion balance determination of Henry Cavendish [1] is well known. The principle is shown in figure 1. Two masses, m are attached to the ends of the horizontal beam (of length $2l$) of a torsion balance, the suspension being a very delicate fibre. Large masses, M, are placed near the masses m.

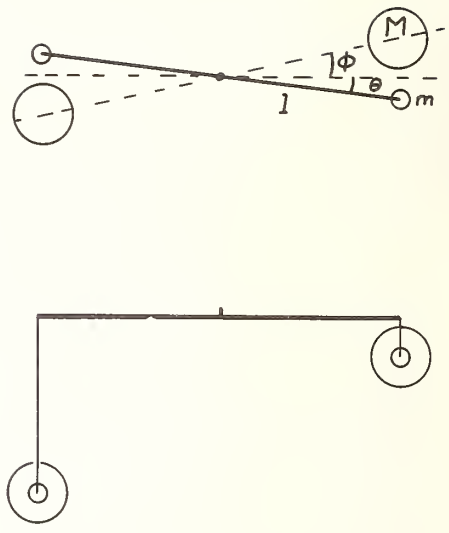


FIGURE 1. *Determination of G with the torsion balance—deflection method.*
m, small mass or beam; M: large attracting mass; l, length of beam; ϕ: angular position of large mass; θ, deflection of beam.

For simplicity, let each mass M be at the same distance l from the point of suspension of the beam. The potential of either large mass at the small mass nearer to it is

$$-GM/l(\phi-\theta),$$

where the angles ϕ and θ are shown as in figure 1, and the couple exerted on the beam is therefore to the first order

$$2(\partial/\partial\theta)\{-GMm/l(\phi-\theta)\}=-2GMm/l(\phi-\theta)^2. \tag{2.1}$$

If the couple exerted by the suspension wire is $\tau\theta$ for a deflection of θ, the equation of motion of the beam is

$$I\ddot{\theta}+\tau\theta-[2GMm/l(\phi-\theta)^2]=0 \tag{2.2}$$

or, supposing that $\theta\ll\phi$,

$$I\ddot{\theta}+\{\tau-(4GMm/l\phi^3)\}\theta=2GMm/l\phi^2 \tag{2.3}$$

The deflection, θ_e, in the equilibrium position is

$$2GMm/\tau l\phi^2$$

while the speed of the oscillation is

$$\{1-(2\theta_e/\phi)\}^{1/2}\omega,$$

where ω is the speed of the free oscillations, namely

$$(\tau/I)^{1/2}.$$

Neglecting the mass of the beam,

$$I=2ml^2$$

and

$$\tau=2ml^2\omega^2. \tag{2.4}$$

Thus

$$\theta_1/G=M/l^3\omega^2\phi^2; \tag{2.5}$$

if all quantities are scaled in the same linear ratio, M is proportional to l^3 and θ_e/G is unchanged by scaling, a result first noticed by Cornu (see [5]). Boys [5] pointed out that if only the length of the torsion beam were reduced, while the period and the large masses were kept the same, the sensitivity would be improved and he conducted an experiment upon that principle. The limitations to the reduction of the length of the beam are: obtaining a fine enough suspension which will still support the beam; placing the large masses sufficiently close to the beam; minimising the attraction of a large mass upon the distant small mass and upon the beam structure.

Boys introduced fibres of fused silica, used ever since for very delicate suspensions, and reduced the two geometrical difficulties by placing the masses at different levels as shown in figure 1. His large masses were 7.4 kg and his small masses were 2.65 g.

He worked with his apparatus in air but went to great pains to minimise convection currents and worked at night so as to avoid mechanical disturbances as much as possible. Braun [6] who used a similar, but larger torsion balance, evacuated his apparatus; his large masses were 9 kg and his small masses 5.4 g.

Nowadays there would not be too much difficulty in measuring masses and distances to about one part in a million even for apparatus with dimension of the order of 10 cm, but the limits to this experiment are set by the instability of the zero on account of creep of the suspension wire, and by mechanical noise.

2.2. The Common Balance

The method was introduced by von Jolly [7] who measured the deflection of a common balance when a large mass was placed below one of the pans, and was considerably developed by Poynting [8]. The principle is shown in figure 2. Masses m and $m+\delta m$ are suspended from the beam of a balance of length $2L$, while a mass M is placed at a distance D below the mean height of the suspended masses. If θ is the angle of deflection of the beam, the condition for equilibrium is

$$g\delta mL\cos\theta+(dM_B-2hm)g\sin\theta$$
$$=[GMm(L\cos\theta+h\sin\theta)/(D-L\sin\theta)^2], \quad (2.6)$$

where M_B is the mass of the balance beam and d is the distance of the centre of mass of the beam below the centre of rotation.

If θ is small,

$$g\delta m=GMm/D^2. \quad (2.7)$$

Alternatively, if δm is zero,

$$\theta=GMmL/gD^2(dM_B-2hm); \quad (2.8)$$

the factor $L/(dM_B-2hm)$ can be found from the sensitivity of the balance when no attracting mass is present:

$$\theta=L\delta m/(2hm-dM_B). \quad (2.9)$$

As is well known, a high sensitivity implies a long period of oscillation of the balance.

Very small forces and deflexions are involved in experiments of this sort.

Poynting, who worked at Birmingham University (England), constructed a large balance to carry a load of some 50 kg with an arm of length about 1 m.

To amplify the angular displacement of the pointer of the balance, he hung a mirror between it and a fixed support (in a way suggested by Lord Kelvin) so that the angular rotation of the mirror was about 150 times that of the pointer.

Poynting's balance, now at the National Physical Laboratory, Teddington, England, was too big to put in an evacuated enclosure, but Poynting placed it in a sealed basement room and operated all controls remotely.

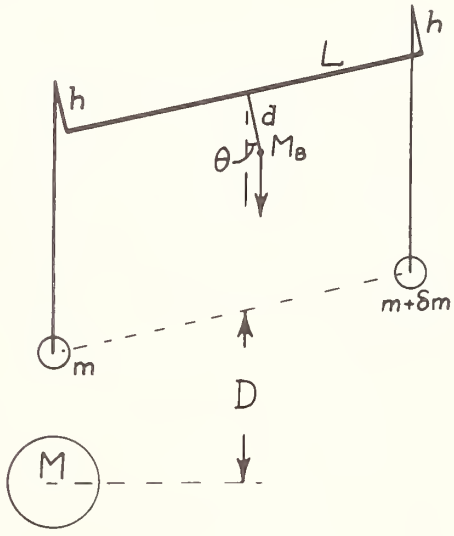


FIGURE 2. *Determination of G with the common balance.*
L, length of arm of beam; *d*, distance of centre of mass of beam below centre of rotation; *h*, height of pan attached above centre of rotation.

The attraction of the large mass upon the beam structure is difficult to calculate, and in addition, the attraction on the mass $(m+\delta m)$ must be allowed for and Poynting moved the mass M to different positions to do so.

2.3. Method Using the Period of a Torsion Balance

The methods of Cavendish and of Poynting both depend on measuring the small change of the position of equilibrium of an oscillating system brought about by the change in attraction of a large mass. The small change has to be measured in the presence of mechanical noise and drift of zero. Generally speaking, the period of an oscillating system can be measured with higher relative precision in such circumstances than can the deflection, and Heyl [9], following pioneer work of Braun [6],

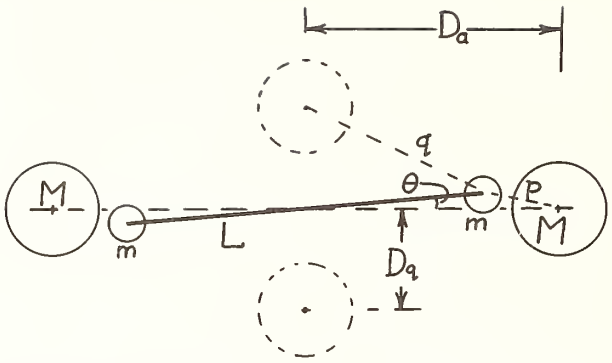


FIGURE 3. *Determination of* G *with the torsion balance—method of periods.*

m, small masses; *M*, large masses (full circles—in line positions, broken circles—quadrature positions); D_a, distance of large mass from centre of suspension, in-line positions; D_q, distance of large mass from centre of suspension, quadrature position, *p*, distance between large and small masses, in-line position; *q*, distance between large and small masses, quadrature position.

who used both the period and the deflection method, carried out a notable series of determinations at the National Bureau of Standards, the latter part of the work in collaboration with Chrzanowski [10].

It has already been seen (sec. 2.1) that the period of a torsion balance is affected by attracting masses. If the masses are placed in line with the beam, the period of oscillation is decreased, whilst if they are placed on a line perpendicular to the beam (quadrature position, see fig. 3) the period is increased. Let the frequencies in the two positions be ω_1 and ω_2 respectively. Then

$$\omega_1^2 = \{\tau + G(A_1 Mm + B_1)\}/I \qquad (2.10a)$$

and

$$\omega_2^2 = \{\tau - G(A_2 Mm + B_2)\}/I, \qquad (2.10b)$$

where τ is the torsional constant of the suspension.

B_1 and B_2 are the attractions of the large masses upon the beam structure in the two positions.

I is the moment of inertia of the beam.

A_1 and A_2 are geometrical factors giving the attraction of the large masses upon the small ones.

In the simple case where the masses m and M may be treated as points, the mutual potential in the in-line position is

$$-2GMm/p = -2GMm/\{L^2 + D_a^2 - 2LD_a\cos\theta\}^{1/2},$$

assuming that $D_a - L$ is much less than L.

The couple, $-\partial V/\partial\theta$, is

$$2GMmLD_a\cos\theta/(L^2 + D_a^2 - 2LD_a\cos\theta)^{3/2},$$

or if $\theta \gg 1$,

$$2GMmLD_a\sin\theta/(D_a - L)^3.$$

Thus A_1 is equal to $2LD_a/(D_a - L)^3$ with the foregoing approximations.

Similarly, the mutual potential in the quadrature position is

$$-2GMm(1/q_+ + 1/q_-),$$

where

$$q_- = (L^2 + D_q^2 - 2LD_q\sin\theta)^{1/2},$$

and

$$q_+ = (L^2 + D_q^2 + 2LD_q\sin\theta)^{1/2}.$$

Then when $\theta \ll 1$,

$$A_2 = 12LD_q/(L^2 + D_q^2)^{5/2}.$$

From eqs (2.10a) and (2.10b)

$$G = I(\omega_1^2 - \omega_2^2)/\{(A_1 + A_2)Mm + B_1 + B_2\}. \qquad (2.11)$$

A_1 and A_2 are factors that involve only the geometrical relations of the masses m and M and are straightforward to calculate, but I, B_1 and B_2 involve integrations over a relatively complicated beam structure and so are difficult to obtain accurately. It will be shown in section 5 how they may be eliminated, but Heyl calculated them.

Heyl's work, though carried out with great attention to detail, is unsatisfactory in other ways. Thus, he measured ω_1 and ω_2 by timing passages of the pendulum through the position of equilibrium and allowing for damping. That would be satisfactory if the oscillations were damped simple harmonic oscillations, but in fact, the equation of motion is highly nonlinear, and Heyl's method of measuring the period may, in consequence, be subject to systematic error (see [10]).

For various, not very fundamental, reasons Heyl used small masses (50 g) of gold, platinum and glass, while the large masses (66 kg) were cylinders of carbon steel. The values of G obtained for the different materials were notably different (table 1) and in view of the high accuracy with which it is known that G is independent of material, there must have been some serious systematic error in Heyl's work. Heyl and Chrzanowski's later work is no more satisfactory.

In particular, the value of G depended on whether the torsion fibre was of hard drawn or annealed tungsten (table 1). Heyl and Chrzanowski considered that they had attained the highest precision of which the torsion balance is capable and in view of the careful study they made of ways of improving the earlier work of Heyl [9], their judgement is probably correct so far as laboratory experiments are concerned. The main difficulties seem to stem from the fact that the torsional couple dominates the gravitational couple when the suspension wire is only a metre or so long. If the suspension can be made much longer (see sec. 5) the gravitational couple may even exceed the torsional couple and then a greater precision may become possible.

2.4. Method of Zahradníček

The principle of Zahradníček's method [11] is shown in figure 4. A torsional pendulum of great inertia (masses M) is set oscillating, and the attraction it exerts on the smaller masses m of a second balance, forces the latter into oscillation. The periods are adjusted to be equal and the ratio of the ampli-

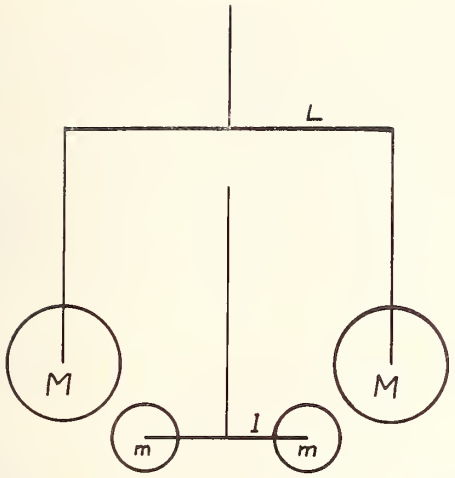


FIGURE 4. *Determination of G with resonant torsion balances.*

tude of the lighter pendulum to the heavier is then proportional to G. An outline of the theory is as follows.

Let a be the length of arm of the lighter pendulum and b be that of the heavier. Let ϕ be the angle of displacement of the heavier pendulum from the common direction of equilibrium and θ that of the lighter. The mutual potential energy of the small masses in the fields of the nearer large masses is then

$$V=-2GMm/\{a^2+b^2-2ab\cos(\phi-\theta)\} \quad (2.12)$$

and the couple exerted on the lighter pendulum, $-\partial V/\partial\Theta$, is therefore

$$-2GMmab\sin(\phi-\theta)/\{a^2+b^2-2ab\cos(\phi-\theta)\}^{3/2}.$$

Let $\tau\theta$ be the couple exerted by the suspension of the lighter pendulum, and suppose that both θ and ϕ are small. The equation of motion of the smaller pendulum is then

$$I\ddot{\theta}+k\dot{\theta}+\left(\tau-\frac{2GMmab}{(a-b)^3}\right)\theta=-\frac{2GMmab}{(a-b)^3}\phi, \quad (2.13)$$

or

$$\ddot{\theta}+\kappa\theta+\omega_0^2\theta=-2GMmab\phi/I(a-b)^3 \quad (2.14)$$

where

$$\kappa=k/I$$

and

$$\omega_0^2=\left[\tau-\frac{2GMmab}{(a-b)^3}\right]I^{-1}.$$

Suppose that $\phi=\phi_0\cos\omega t$, the damping of the larger pendulum being supposed negligible. Then

$$\theta=\frac{\phi_0}{\{(\omega_0^2-\omega^2)^2+\alpha^2\omega^2\}^{1/2}}\times\frac{2GMmab}{I(a-b)^3}\cos(\omega t+\chi) \quad (2.15)$$

where

$$\tan\chi=-\kappa\omega/(\omega_0^2-\omega^2).$$

The maximum amplitude of θ is attained when

$$\omega_1^2-\omega^2=\tfrac{1}{2}\kappa^2$$

and is

$$2GMmab\phi_0/I\kappa\omega(a-b)^3$$

if $\kappa\ll\omega$.

It will be seen that the determination of G requires accurate measurements of ϕ_0, of κ, and of the moment of inertia of the small pendulum, as well as a geometrical factor which, in detail, is more complex than $ab/(a-b)^3$.

If $\kappa\ll\omega$, the lighter beam oscillates with a phase shift of nearly $\frac{1}{2}\pi$ from the heavier.

Zahradníček's result is given in table 1.

3. Difficulties and Sources of Error

There are four important sources of error common to all the foregoing methods, namely mechanical noise; the determination of the centres of mass of attracting bodies; the attraction of beams, supports and so on; and variability of the suspension.

There is little doubt that the limits of accuracy of all existing determinations are set by mechanical disturbance of the oscillating system and by creep in suspensions. In every deflection method the steady displacement has to be measured in the face of appreciable oscillation and drift and it was the difficulty of such observations of deflections that led Heyl and Zahradníček to employ oscillating systems. Zahradníček's method is of course no improvement, since an amplitude has still to be observed, and his result is no better than others. Heyl's method would have been expected to show a considerable improvement in precision, for the measurement of the period of an oscillating system should be much more precise than the measurement of an amplitude, provided it is extended over a sufficiently large number of oscillations.

A torsional pendulum in an evacuated vessel should be least subject to mechanical noise because rotation about a vertical axis is a very minor part of ground movement, but there may be coupling through inelastic relaxation of the suspension fibre. The period of the suspended system should be made as long as possible both to give high sensitivity to the gravitational forces and to minimise the effects of mechanical disturbance.

Suppose the upper anchorage of the torsion fibre undergoes a rotation through an angle $\phi_0\exp(i\omega t)$. The forced motion will have an amplitude θ_0 equal to

$$-\phi_0\omega_0^2/\omega^2$$

if ω_0, the free speed of the torsion balance, is much less than ω. ω_0 should therefore be made as small as possible.

Again, suppose that a couple $g_0\exp(i\omega t)$ is exerted on the beam (for example, due to motions of any

ambient air). The amplitude of the forced motion is then

$$-g_0/I\omega^2,$$

where I is the moment of inertia of the suspended system. I should therefore be as large as possible.

The second general problem is the determination of the centres of mass of the various masses.

Small spheres, up to say 10 cm in diameter, can be formed of glass or hard metal with very high accuracy, and by rotating them about the centres of figures between different sets of observations, the effect of any departure of the centre of mass from the centre of figure may be eliminated. If the masses exceed a few kilogrammes handling difficulties prevent their being made spherical with high accuracy; it is not possible to determine the centre of mass experimentally with adequate precision (also because of difficulties of handling) and it must then be assumed that the centre of mass coincides with the centre of figure—which can be found with high precision by conventional metrology. The two centres will not coincide if the material is not uniform or has cavities within it.

There is then much to be said for making the large masses in sections so that the uniformity of each section can be more readily checked. The material must be chosen so that it may be worked or cast in vacuum to avoid holes.

The structure of the beam of the torsion balance or other oscillating system is usually not of a simple form and it has in the past been necessary to calculate its moment of inertia and the attraction of large masses upon it in great detail, a very troublesome undertaking.

Heyl and Chrzanowski tested a differential method, altering the small masses on the torsion beam. Their results were not satisfactory, probably because the stress on the torsion fibre is changed. It will be shown in section 5 that a differential procedure may be used in which the stress is kept constant. It is plain from these experiments of Heyl and Chrzanowski, that the behaviour of a torsion fibre depends on the stress to which it is subject and on the stress history, and the divergent results they obtained with hard drawn and annealed tungsten (table 1) point to the same conclusion. It is clear that if a torsional suspension is to be used, it should not exert a torque large compared with the gravitational torque; the consequence of this consideration is that the suspension should be as long and the masses as heavy as possible (sec. 5).

No particular care has been taken in the past with magnetic forces, although non-magnetic materials have usually been employed and Heyl employed soft iron screening. Elementary calculations show that the magnetic forces between ordinary "non-magnetic" materials are negligible in comparison with the gravitational forces, in part because the sources are dipoles.

An important consideration in the experiments of Heyl and of Zahradníček is that the equations of motion of the oscillating system may be highly non-linear. The reason is that the large and small masses of a pair are placed as close together as possible to give a large gravitational couple and in consequence their relative positions change greatly when the beam rotates. The couple thus depends appreciably upon the cube and higher odd powers of the angle of displacement of the beam. The effects of a non-linear motion have not hitherto been explored in detail but certainly must be considered when deciding just what is done in measuring a "period" of oscillation.

4. The Determination of J. W. Beams and His Collaborators

Beams and his colleagues [12, 13] have developed a new approach to the determination of G in which

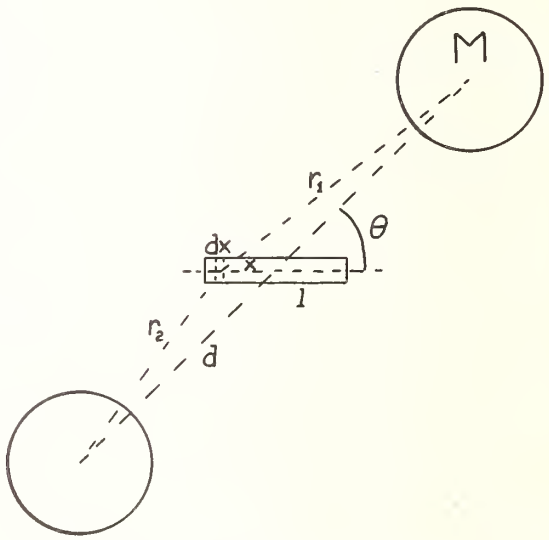


FIGURE 5. *Determination of G by direct measurement of acceleration.*

they measure the acceleration of a suspended system under the attraction of heavy masses. The principle is shown in figure 5.

A horizontal rod of length l is hung from a torsion fibre of fused silica attached to the mid-point and is attracted by masses M in the horizontal plane containing the rod. The masses together with the anchorage of the torsion fibre are carried on a table that can rotate about the vertical axis containing the torsion fibre. The attraction of the masses causes the rod to turn towards them, decreasing the angle θ. The angle of the rod relative to the table is measured by an optical lever system and the table is rotated by a servo system to keep that angle (and therefore θ) a constant.

Thus the acceleration of the table is equal to the acceleration of the rod under the attraction of the large masses at the angle θ.

Consider the mutual potential of the masses M and the element of rod of length δx at x from the centre. If the mass per unit length is μ, the mutual potential, δV, is

$$-GM\mu\delta x\{r_1^{-1}+r_2^{-1}\}.$$

The torque on the rod arising from this mutual potential is

$$\delta C = -\partial\delta V/\partial\theta$$

$$= GM\mu xd\sin\theta\left[\frac{1}{(x^2+d^2+2xd\cos\theta)^{3/2}} + \frac{1}{(x^2+d^2-2xd\cos\theta)^{3/2}}\right]\delta x. \quad (4.1)$$

The total couple, C, on the rod is

$$2\int_0^l \delta C$$

and may be written as

$$(2GM\mu\sin\theta/d)f(\theta,\lambda),$$

where $\lambda = l/d$, $\xi = x/d$ and

$$f(\theta,\lambda) = \int_0^\lambda \xi d\xi\left[\frac{1}{(1+\xi^2+2\xi\cos\theta)^{3/2}} - \frac{1}{(1+\xi^2-2\xi\cos\theta)^{3/2}}\right]. \quad (4.2)$$

If $\lambda \ll 1$, as is the case in practice,

$$f(\theta,\lambda) = 3\lambda^2\cos\theta \quad (4.3)$$

and so

$$C = (3GM\mu\lambda^2\sin 2\theta/d), \quad (4.4)$$

which has a maximum for $\theta = 1/4\pi$. The angular acceleration, $\ddot{\theta}$, is equal to C/I, where I, the moment of inertia of the rod about the vertical axis, is

$$\tfrac{2}{3}\mu d^3\lambda^3.$$

Hence

$$\ddot{\theta} = \tfrac{9}{2}(GM\sin 2\theta/\lambda d^3) \quad (4.5)$$

and the largest acceleration is obtained by making λ and d small.

Further details are given by Towler and others [12]. The preliminary experiments so far carried out have used a quartz fibre suspension but it is intended to use a magnetic suspension.

5. The New Determination at Trieste

A new determination by the Braun-Heyl method (section 2.3) has for some time been in preparation by Professor A. Marussi and myself [14]. It is a cooperative project of the Instituto di Geodesia e Geofisica of the University of Trieste, of the National Physical Laboratory, Great Britain and the Department of Geophysics, University of Edinburgh. For reasons given below the experiment is being set up in the Grotta Gigante, a large limestone cave near Trieste.

The new determination differs from that of Heyl in a number of important ways.

In the first place, Heyl pointed out that there is considerable advantage in using a very large system (see also [14]). Boys himself [5] appreciated that even in the deflection mode of the torsion pendulum, reduction in size is not necessarily an advantage, even though the period is kept long; in the oscillation mode of the torsion pendulum, somewhat different considerations are involved [14] and the advantage of a large system seems clear. The difficulties are essentially the practical ones of handling a large system and adjusting and observing a system with a very long periodic time.

The Grotta Gigante gives a clear height of 80 m for the torsional suspension and it is possible for the first time to obtain a torsional couple less than the gravitational couple exerted by the large masses. In fact, such a weak suspension is not used because the system would be unstable with the large masses in the quadrature position, but it does mean that the periods of oscillation can be chosen at will within a wide range. The aim will be to choose periods that give a maximum difference $(\omega_1^2-\omega_2^2)$ and which are at the same time far from natural periods of earth motion in the cave. The pendulum is in an evacuated enclosure, which limits the length of the beam to 1 m. The small masses will be 10 kg. In the in-line position the large masses will be 2 m apart, in the quadrature position, 1.04 m.

The second important feature is that the experiment will be done differentially to eliminate both the moment of inertia and the gravitational attraction of the beam.

Let I_0 be the moment of inertia of the beam without the small masses. Let the separation of the small masses be $2d$. Then the total moment of inertia is

$$I_0+2m(d^2+k^2)$$

where k is the radius of gyration of a small mass. Hence (see section 2.3)

$$G(A_1Mm+B_1)+\tau = \omega_1^2\{I_0+2m(d^2+k^2)\} \quad (5.1a)$$

in the in-line position, and

$$-G(A_2Mm+B_2)+\tau = \omega_2^2\{I_0+2m(d^2+k^2)\} \quad (5.1b)$$

in the quadrature position. Now let the small masses be moved in to positions in which they are separated by $2d'$ $(d'<d)$. Then there are two corresponding

equations

$$G(A_1'Mm+B_1)+\tau=\omega_1'^2\{I_0+2m(d'^2+k^2)\}, \quad (5.2a)$$

$$-G(A_2'Mm+B_2)+\tau=\omega_2'^2\{I_0+2m(d'^2+k^2)\}, \quad (5.2b)$$

where, as before A_1' and A_2' are geometrical factors depending on the positions of the centres of mass of the large and small masses. B_1 and B_2 are unaltered because the positions of the large masses are unaltered. B_1, B_2, and I_0 may then be eliminated.

$$GM\frac{A_1-A_1'}{\omega_1^2-\omega_1'^2}+\frac{A_2-A_2'}{\omega_2^2-\omega_2'^2} = \frac{2(\omega_1'^2\omega_2^2-\omega_1^2\omega_2'^2)(d^2-d'^2)}{(\omega_1^2-\omega_1'^2)(\omega_2^2-\omega_2'^2)}. \quad (5.3)$$

It will be seen also that neither the mass nor radius of gyration of a small mass is required.

Heyl made the large masses in the form of cylinders. We have observed that by choosing cylinders of the correct ratio of length to diameter, and providing them with additional end caps, it is possible to eliminate the zonal spherical harmonic terms of order 2, 4, and 6 in the gravitational potential, and thus the potential is that of a point mass at the centre of figure, together with small correcting terms that vary as $1/r^9$ or faster. Such cylinders can be made of discs and each will in fact be a pile of discs of total mass 500 kg. In this way, a very large mass can be constructed, yet the homogeneity can be checked by determining the centre of mass of each component disc; the discs will be assembled to place the centre of mass as close as possible to the centre of figure of the assembly. Large and small masses are made of a copper-aluminum alloy of low magnetic susceptibility.

As has been said before, the motion of the pendulum is non-linear and follows an equation

$$I\ddot{\theta}+a\dot{\theta}+b_1\theta+b_3\theta^3+b_5\theta^5\cdots=\epsilon. \quad (5.4)$$

If a, b_3, b_5, were zero, the period $(b_1/I)^{1/2}$ could be found from a power spectrum analysis of the position of the pendulum, but when a, b_3, b_5, cannot be neglected, that analysis is no longer possible. The procedure being adopted is to generate in a computer a position, θ_t, using trial value of the coefficients in the differential equation, to form the difference $(\theta-\theta_t)$ and from that to derive improved values of the parameters, especially of b_1/I.

The position of the pendulum, Θ, is accordingly to be measured at intervals of about 20 s (for periods of some 2000 s). The measurements are to be made with a simple two-beam laser interferometer, the output of which will be in the form of counts of fringes.

It will be seen that the experiment is quite classical in principle but by improvements in design and the use of new methods of data recording and reduction, it is hoped that a greatly improved precision will be attained. The highly stable temperature in the Grotta Gigante (a range of less than 1 °C in a year) should also help. The measurements of mass, length and time will be made to better than 1 part in 10^5 and the limit on accuracy will undoubtedly be mechanical noise, the nature of which is at present unknown, although the periods of certain disturbances, such as tides and oscillations in the Gulf of Venice are known.

6. Conclusion

No method so far undertaken or proposed comes near to giving the accuracy needed to investigate changes of G with direction or in time, and it is premature to say whether current work will lead to any new ideas on that problem. It is certainly too soon, with the large uncertainties indicated in table 1, to attempt any critical assessment of the current value of G.

It may perhaps be mentioned that delicate experiments have shown there to be no gravitational shielding by one body between two others. It is possible to infer this result from the high regularity of the orbits of the Moon and the planets, but a direct demonstration has been given by Professor A. Marussi.

One of the instruments in the Grotta Gigante at Trieste is a very sensitive horizontal pendulum (pivoted about a nearly vertical axis) for measurements of tidal variations of the direction of gravity. The 1964 solar eclipse occurred with the Sun and the Moon in line due west of Trieste, so that if the Moon shielded the gravitational attraction of the Sun on the pendulum bob, a discontinuity in the position of the pendulum would have occurred at eclipse. None such was detected and very severe limits can be placed on any shielding effect.

7. References

[1] Cavendish, H., Philos. Trans. Roy. Soc. **88**, 467 (1798).
[2] Eötvös, R. V., Pekar, D., and Fekete, E., Ann. Physik **68**, 11(1922).
[3] Roll, P. G., Krotkov, R., and Dicke, R. H., Ann. Phys. (USA), **26**, 442 (1964).
[4] Sakuma, A., Proceedings of this conference.
[5] Boys, C. V., Philos. Trans. Roy. Soc. A, **186**, 1 (1895).
[6] Braun, K., Denkshr. K. Akad. Wiss. Wien, math. naturw. Kl. **64**, 187 (1896).
[7] von Jolly, Abhand Bayr. Akad. 2Kl 13 and 14 (1881).
[8] Poynting, J. H., Philos. Trans. Roy. Soc. A, **182**, 865 (1891).
[9] Heyl, P. R., J. Res. Nat. Bur. Stds. **5**, 1243 (1930).
[10] Heyl, P. R., and Chrzanowski, P., J. Res. Nat. Bur. Stds. **29**, 1, (1942).
[11] Zahradníček, J., Phys. Zeit. **34**, 126, (1933).
[12] Rose, R. D., Parker, H. M., Lowry, R. A., Kuhlthau, A. R., and Beams, J. W., Phys. Rev. Lett. **23**, 655 (1969).
[13] Towler, W. R., Parker, H. M., Lowry, R. A., Kuhlthau, A. R., and Beams, J. W., Proceedings of this conference.
[14] Cook, A. H., Contemp. Phys. **9**, 227 (1968).

D. Halford: Do you plan to servo the angular position of the torsion element where it is attached up above to make a null experiment?

A. H. Cook: No, we do not. One could I suppose do this but my feeling is that one should have as simple an experiment as possible, at least to begin with. Maybe we'll have a go later.

H. Metcalf: You haven't told us how accurately you expect to be able to get *G*.

A. H. Cook: I was hoping no one would ask. (*Laughter*) I said how accurately we would measure the components, the mechanical components, approaching 1 in 10^6 I think. It's really just guessing to say what sort of accuracy we shall get out of the measurement. It does depend on what the mechanical noise is, and I suspect our experiments are going to be the most sensitive detectors of mechanical noise. We don't know. We just hope we have made the right choice in going to this site and doing it the way we are. And there's the problem of irregularities in suspensions which, of course, we can do something about by changing them maybe. I naturally hope that we shall do better than Cavendish. (*Laughter*) I hope we will do two or three orders of magnitude better than Cavendish, but it's too early to say.

H. Preston-Thomas: I thought you were just going to leave it without even the additional orders.

E. R. Cohen: Isn't it true that going to small size doesn't really gain you that much because the masses also have to scale as the cube of the dimensions, and only to the extent that you can use relatively larger masses do you really gain by going to smaller size?

A. H. Cook: This is right. This famous argument of Boys has got into all the textbooks, but I think very much oversimplified, because if you read Boys' original paper you will find that he himself pointed out there were severe limits on its applicability of the sort that you mentioned. And, as I say, when you go to a different type of experiment from Boys', then it's not applicable.

B. P. Kibble: Is it possible that you might have gravitational gradients in the horizontal plane in your cave which might affect your result?

A. H. Cook: That is right; but in fact, it doesn't matter to the first order. But we shall make a survey of the gravity field to see what the gradients are like. This is a fairly trivial thing to do. It's only the change of gravitational potential with angle that matters, of course, so provided there isn't a large change of gravity horizontally in the direction in which the beam is set, you're all right. If you make a gravity survey, then you can set the beam to the optimum position; but I don't think this is likely to be a serious matter.

H. Preston-Thomas: If I can make one final remark, I notice that you described the force in terms of milligrams, and so I detect a residual interest in *g*. (*Laughter*).

Measurement of the Newton Gravitational Constant G

W. R. Towler, H. M. Parker, R. A. Lowry, and A. R. Kuhlthau

Department of Aerospace Engineering and Engineering Physics, University of Virginia Charlottesville, Virginia 22901

and

J. W. Beams

Department of Physics, University of Virginia, Charlottesville, Va. 22901

A novel method for measuring the gravitational constant G utilizes a modified Cavendish arrangement in which the gravitational interaction between a large mass system and a small mass system produces a torque on the latter. A rotary table carrying the large masses is servoed to constant relative geometry with respect to the small mass system. The angular acceleration of the table and the other appropriate parameters are measured. Recently [Phys. Rev. Letters **23**, 655 (1969)] a slightly improved value of G was reported. Further improvements in the method and noise reduction are described and the ultimate potential accuracy of the method is discussed.

Key words: Gravitation; Newton gravitational constant; universal gravitational constant.

1. Introduction

Gravitational interaction possesses two unique properties; its universality and its extreme weakness. Both of these features contribute to the difficulty of measuring the gravitational constant G which occurs in Newton's law of gravitation

$$F=G(m_1m_2/d^2) \qquad (1)$$

where F is the force of attraction between any two particles of matter in the universe having masses of m_1 and m_2, and d is the distance between the particles. The relatively small size of masses which can be used in practical laboratory experiments to determine the gravitational constant, essentially requires the absolute measurements of very minute forces or torques. Furthermore, these forces or torques must be measured in the presence of disturbing forces and force gradients which result from the asymmetry of the mass distribution around the experiment, i.e., the test masses interact not only with each other but also with every other mass in the universe, for it has not been possible to isolate or shield them from these perturbing forces.

Critical reviews [1, 2] have been given of a number of most ingenious experiments devised and carried out by highly talented and skillful workers for measuring G. However, it is evident that there still exists an urgent need for determining with as much accuracy as possible both the absolute value of G and a possible variation of G with time and other factors.

A few years ago the authors [3] proposed a new experimental method for determining G and some preliminary results have been briefly reported [4, 5]. This paper will present additional preliminary results as well as a more detailed description of the method and the apparatus used. The principle of the method is illustrated in figure 1.

Two large spherical masses (tungsten spheres) are mounted on a rotary table which can be driven about its axis of rotation by a specially designed electric motor. Also mounted from the same rotary table is a gas-tight chamber in which a small horizontal cylinder is suspended by means of a quartz torsion fiber fastened to the top of the cylinder and hanging in the axis of the chamber. This small horizontal cylinder is commonly called the small mass system as contrasted to the large spheres which are referred to as the large mass system.

The gravitational interaction between the two mass systems tends to cause the small mass system to deflect so as to bring its axial center line in alignment with the line connecting the centers of the two large spheres. This changes the angle, θ. However, a beam of light from a source mounted on the rotary table is reflected from a mirror mounted on the small mass system near the axis of the quartz fiber and falls on a photodiode, also mounted on the rotary table. Thus the light beam generates an angle β.

As θ begins to change due to the gravitational interaction, so does β. The photodiode senses very minute changes in β, and this "error" signal is used

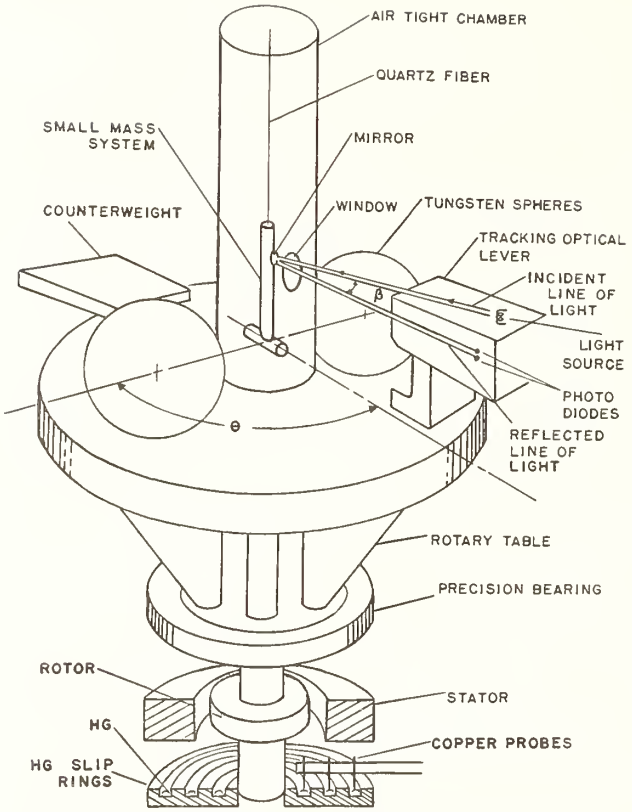


FIGURE 1. *Schematic drawing of experimental apparatus.*

to drive the motor which rotates the table so as to maintain β (and hence θ) constant. Thus with the angular separation, θ, remaining constant, the small mass system experiences a constant torque, which in turn causes a constant angular acceleration of the rotary table. This acceleration can be determined very accurately by measuring the period of the rotating table, and will be shown to be a direct measure of G.

This method possesses two novel features which contribute to the potential for improved accuracy. First, the interaction force of the two masses is manifested in an acceleration (change of rotational velocity) rather than a deflection. The effect of the interaction is cumulative and can be integrated over a long period of time, thus improving precision. Second, the two mass systems rotate about an axis many times during a measurement, and hence the effects of gravitational fields or field gradients due to extraneous masses are effectively cancelled except for higher order effects.

2. Theoretical Analysis of the Gravitational Torque

The analysis of the gravitational torque on the small mass system in figure 1 exerted by the spheres can be separated into two distinct parts. In the first part the gravitational potential at a point outside the cylinder is calculated and in the second part the potential gradient is used to determine the torque and acceleration the small mass system will experience.

Referring to figure 2, the gravitational potential ϕ on the z axis for $z>C$ of a ring of mass m, radius r and located at $z=C\cos\alpha$ is

$$\phi_{\text{ring}}(P') = \frac{mG}{(R^2+C^2-2RC\cos\alpha)^{1/2}}. \tag{2}$$

This can be expanded to Legendre polynomials to give

$$\phi_{\text{ring}}(P') = mG\sum_{l=0}^{\infty}\frac{C^l}{R^{l+1}}P_l(\cos\alpha). \tag{3}$$

The gravitational potential at any point in space is obtained by multiplying eaeh member of this series by $P_l(\cos\theta)$.

$$\phi_{\text{ring}}(R,\theta) = mG\sum_{l=0}^{\infty}\frac{C^l}{R^{l+1}}P_l(\cos\alpha)P_l(\cos\theta). \tag{4}$$

The potential for a cylinder is obtained by writing this equation in differential form and integrating over r and z.

$$\phi(R,\theta) = \frac{2\pi\rho G}{R}\sum_{l=0}^{\infty}\frac{P_l(\cos\theta)}{R^l} \times\int_{-L/2}^{+L/2}\int_0^a C^l P_l(\cos\alpha)\,r\,dr\,dz \tag{5}$$

where ρ is the density of the cylinder.

Performing the integrations in (5) using Rodrigues' formula for $P_l(\cos\alpha)$ and expanding it by means of the binomial expansion yields,

$$\phi(R,\theta) = \frac{mG}{R}\sum_{l=0}^{\infty}\sum_{k=0}^{l/2}\sum_{j=1}^{k+1}A_{jkl}\left(\frac{2a}{L}\right)^{2(j-1)}\left(\frac{L}{2R}\right)^l P_l(\cos\theta) \tag{6}$$

where $m=\rho\pi a^2 L$,

$$A_{jkl} = \frac{A_{kl}}{(k+1)(l-2j+3)}\binom{k+1}{j}, \tag{7}$$

$$A_{kl} = \frac{(-1)^k}{2^l l!}\binom{l}{k}\frac{(2l-2k)!}{(l-2k)!}, \tag{8}$$

$$\binom{l}{k} = \frac{l!}{k!(l-k)!}, \tag{9}$$

and a, R, and L are as shown in figure 2. The potential is nonvanishing only for even values of l.

This potential may now be used to determine the magnitude of the force on a point mass, M, located at the point P,

$$\bar{F} = M\nabla\phi. \tag{10}$$

This is the same force that M exerts on the cylindrical mass by Newton's third law. The component of force perpendicular to R is the only one

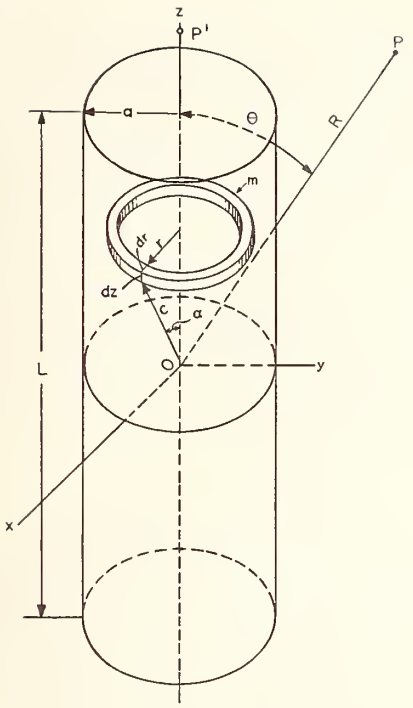


FIGURE 2. *Geometry for discussion of gravitational torque on a cylinder.*

which gives rise to a torque about 0. In spherical coordinates, this component is given by

$$\bar{F}_\theta = (M/R)\,(\partial\phi/\partial\theta)\hat{\theta}. \tag{11}$$

When two equal point masses are placed diametrically opposite each other at (R, θ) and $(R, \theta+\pi)$ a pure torque about the x-axis results, given by

$$T = 2M(\partial\phi/\partial\theta). \tag{12}$$

Applying this equation to eq (6) yields

$$T = \frac{2mMG}{R}\sum_{l=0}^{\infty}\sum_{k=0}^{l/2}\sum_{j=1}^{k+1} A_{jkl}\left(\frac{2a}{L}\right)^{2(j-1)} \times\left(\frac{L}{2R}\right)^{l}\frac{\partial}{\partial\theta}\left[P_l(\cos\theta)\right]. \tag{13}$$

The angular acceleration acting on the small mass system is

$$\dot{\omega} = T/I. \tag{14a}$$

The stem holding the cylindrical mass is assumed axially symmetric about axis y (see fig. 1) and hence will not contribute to T but will contribute to I which is the moment of inertia of the torsional pendulum. With care, the contribution of the mirrors to T can be minimized and represents a negligible source of error. Their contribution to I is readily calculable and was taken into account.

Hence the angular acceleration acting on the torsional pendulum is

$$\dot{\omega} = \frac{2mMG}{IR}\sum_{l=0}^{\infty}\sum_{k=0}^{l/2}\sum_{j=1}^{k+1} A_{jkl}\left(\frac{2a}{L}\right)^{2(j-1)} \times\left(\frac{L}{2R}\right)^{l}\frac{\partial}{\partial\theta}\left[P_l(\cos\theta)\right]. \tag{14b}$$

Evaluating the $l=2$, 4, and 6 poles yield,

$$\dot{\omega}_2 = (3mMG/IR^3)\left[\tfrac{1}{3}-\tfrac{1}{4}(2a/L)^2\right](L/2)^2\sin 2\theta \tag{15}$$

$$\dot{\omega}_4 = \dot{\omega}_2\,\frac{5}{6}\left(\frac{L}{2R}\right)^2\frac{\left[\tfrac{1}{5}-\tfrac{1}{2}(2a/L)^2+\tfrac{1}{8}(2a/L)^4\right]}{\left[\tfrac{1}{3}-\tfrac{1}{4}(2a/L)^2\right]} \times(7\cos^2\theta-3) \tag{16}$$

$$\dot{\omega}_6 = \frac{\dot{\omega}_2}{48}\left(\frac{L}{2R}\right)^4\frac{\left[\tfrac{1}{7}-\tfrac{3}{4}(2a/L)^2+\tfrac{5}{8}(2a/L)^4-\tfrac{5}{6}\tfrac{5}{4}(2a/L)^6\right]}{\left[\tfrac{1}{3}-\tfrac{1}{4}(2a/L)^2\right]} \times[1386\cos^4\theta-1260\cos^2\theta+210]. \tag{17}$$

Of course, $\dot{\omega}$ is

$$\dot{\omega} = \dot{\omega}_2+\dot{\omega}_4+\dot{\omega}_6+\cdots \tag{18}$$

where the higher terms are negligible for the current measurement aceuracy of $\dot{\omega}$.

Equation (18) is correct if the torsional constant, k, of the pendulum is zero. If $k\neq 0$ then the measured angular acceleration, α, is

$$\alpha = \dot{\omega}\pm\dot{\omega}_{w/o} \tag{19}$$

where $\dot{\omega}_{w/o}$ is the angular acceleration without (w/o) the spheres on the rotary table. The plus sign of course applying to $\dot{\omega}$ and $\dot{\omega}_{w/o}$ being in the same direction.

An expression for G is easily obtained. From eq (19),

$$\dot{\omega} = \alpha\pm\dot{\omega}_{w/o}. \tag{20}$$

Using eq (15) through (17), eq (18) may be written as,

$$\dot{\omega} = GA(1+B+C+\cdots) \tag{21}$$

where A, B, and C follow clearly from (15) through (17).

Since I is actually the moment of inertia of the small mass system (the moment of inertia of the fiber is negligible) which consists of the horizontal cylinder and the vertical stem,

$$I = I_r+I_s. \tag{22}$$

This may be written as

$$I = m(L/2)^2\left[\tfrac{1}{3}+\tfrac{1}{4}(2a/L)^2+(2/L)^2(I_s/m)\right]. \tag{23}$$

Hence an expression for G, is

$$G = (\alpha\pm\dot{\omega}_{w/o})/A(1+B+C\cdots) \tag{24}$$

where

$$A = \frac{3M}{R^3}\,\frac{\left[\tfrac{1}{3}-\tfrac{1}{4}(2a/L)^2\right]\sin 2\theta}{\left[\tfrac{1}{3}+\tfrac{1}{4}(2a/L)^2+(2/L)^2(I_s/m)\right]}, \tag{25}$$

$$B = \frac{5}{6}\left(\frac{L}{2R}\right)^2\frac{\left[\tfrac{1}{5}-\tfrac{1}{2}(2a/L)^2+\tfrac{1}{8}(2a/L)^4\right]}{\left[\tfrac{1}{3}-\tfrac{1}{4}(2a/L)^2\right]}(7\cos^2\theta-3), \tag{26}$$

and

$$C=\frac{1}{48}\left(\frac{L}{2R}\right)^4\frac{\left[\frac{1}{7}-\frac{3}{4}(2a/L)^2+\frac{5}{8}(2a/L)^4-\frac{5}{64}(2a/L)^6\right]}{\left[\frac{1}{3}-\frac{1}{4}(2a/L)^2\right]}$$
$$\times[1386\cos^4\theta-1260\cos^2\theta+210]. \quad (27)$$

The experimental problem is to determine α, $\dot{\omega}_{w/o}$, R, M, a, L, I_s, m, and θ as accurately as possible. Then G can be calculated from eq (24). Additional terms can be calculated if necessary, but they do not contribute to the present level of precision.

3. Description of Apparatus

Figure 3 is a photograph of the experimental apparatus which is mounted inside a clean temperature controlled (0.01 °C) room. Two concrete block columns mounted on a concrete base (isolated from the building) support the experimental apparatus. The rotary table (1) and precision air bearing (2) support a gas tight chamber (3) which encloses a torsional pendulum immersed in helium gas to reduce any temperature gradients. The tungsten spheres (4) are placed on their stands so that their radial position relative to the center of the chamber is accurately known. The tracking optical lever (5) is of a type due to Jones and Richards [6], and monitor any angular change of the torsion pendulum inside the chamber. A servo mechanism is designed to maintain a zero output from the sensing photodiodes of the optical lever. To do this it rotates the rotary table (1) by means of an ac torque motor. The rotor of this torque motor is inside the circular stator (6) and is attached to the table through the precision bearing. To the bottom of the rotor is attached the mercury slip ring assembly (7) which provides electrical connections to the tracking optical lever. The water cooled stator (6) is mounted to an aluminum plate and the assembly is held tightly against the Rahn Granite Surface Plate (8) by four clamps (9). This allows for easy and reliable positioning of the stator relative to the rotor. The timing optical lever (10) senses the passage of the small mirror (11) mounted to the rotary table. This allows the period of the table to be calculated. From these periods the angular accelerations α and $\dot{\omega}_{w/o}$ can be calculated. The measurement of the other variables in eq (24) leads to the determination of G.

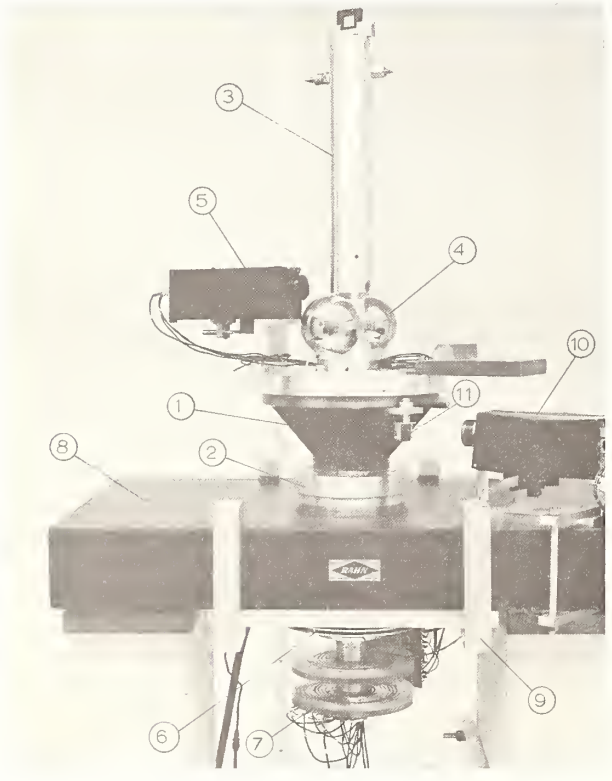


FIGURE 3. *Photograph of apparatus.*

4. Determination of G

In order to obtain a value for G, the nine parameters appearing in eq (24) must be determined. Five of these (a, L, I_s, M, and m) can be considered as constants of the apparatus and can be measured directly, independent of the experimental observations. The other four (α, $\dot{\omega}_{w/o}$, R, and θ) must be determined as a part of the observations.

The high-density tungsten spheres, which make up the large mass system (figs. 1 and 3) were specially made at the Y-12 plant of the Union Carbide Corporation Nuclear Division, Oak Ridge, Tenn. [7]. Each of the tungsten spheres had a mass of approximately 10 Kg and rested on a damped three point mount which in turn was supported by a common large quartz plate in order to minimize the effect of temperature fluctuations. The distance between the centers (fig. 1) is approximately 12 cm and was measured by standard methods using Johansson gauge blocks as references.

The small mass system consists of a carefully machined high purity oxygen free (diamagnetic) copper cylinder approximately 4 cm long (L) and 0.2 cm in diameter (a) fastened to a small accurately machined aluminum alloy rod which in turn was supported by a 25 μ quartz fiber 33 cm long hung from the top of the metal chamber containing helium as shown in figure 1. The mass of the copper cylinder m is about 4 g. The quartz fiber and aluminum alloy stem were carefully positioned in the vertical axis of rotation and the axis of the copper cylinder was in the horizontal plane containing the line between the center of the two tungsten spheres. The aluminum stem carried two small mirrors as shown in figure 1. The angle θ' between the axis of the copper cylinder and perpendicular to the tracking mirror on the aluminum stem is 44° 43′±0.5′.

The dimensions chosen for the small mass system were a compromise between feasible size and keeping the distance between the two spheres $2R$ as small as possible for maximum angular acceleration α. The precise physical dimensions of the mass systems are shown in table 1.

TABLE 1. *Mass systems*

Large masses—high density tungsten spheres		Sphere #1	Sphere #2
Mass	kg	10.489980±0.00007	10.490250±0.00007
Diameter	cm	10.165072	10.165108
Distance between center of mass and geometrical center	cm	4.610×10^{-4}	7.569×10^{-4}
Sphericity	cm	12×10^{-6}	12×10^{-6}
Small mass system—copper cylinder			
Length L	cm	3.9649±0.0004	
Diameter a	cm	0.19824±0.0004	
Mass m	g	4.0512±0.0001	
Moment of inertia of aluminum alloy system	g cm²	0.0408±0.0001	
Angle between the axis of copper cylinder and perpendicular to tracking mirror		44°43′±0.5′	

The angle θ between the longitudinal axis of the cylindrical rod of the small mass system and an imaginary line joining the centers of mass of the spheres not only must be determined but must be held constant during the experiment. As can be seen from eq (15), it is desirable to make this angle approximately 45° in order to maximize the angular acceleration $\dot{\omega}$ of the table. In effect, this will minimize the experimental error since when θ is zero or 90° there is no torque exerted on the small mass system. Consequently the torque must pass through a maximum when θ is varied from zero to 90°. It turns out that there is a relatively flat maximum near $\theta=45°$. In order to maintain θ constant, β is held constant by the servo system shown in the block diagram in figure (4). The details of this servo system are explained in the literature [8]. Experimental tests show that the tracking error of this system is approximately ±0.2 s of arc. The small mass system is suspended with the perpendicular to the tracking mirror also perpendicular to the line joining the centers of mass of the spheres. Consequently θ is very nearly equal to $90°-\theta'$ with small corrections determined from the position of the tracking optical lever [5]. The tracking optieal lever is positioned so that the tracking angle corresponds as closely as possible to the neutral position of the small mass systems with the tungsten spheres removed. This minimizes the background angular acceleration of the table $\dot{\omega}_{w/o}$ due to twist in the quartz fiber.

The measurement of the angular acceleration α and $\dot{\omega}_{w/o}$ in eq (24) involves the determination of the period of the rotary table. Figure (5) is a block diagram of the timing system. When the table turns the attached mirror reflects the light back to the fixed optical lever producing a sharp voltage spike.

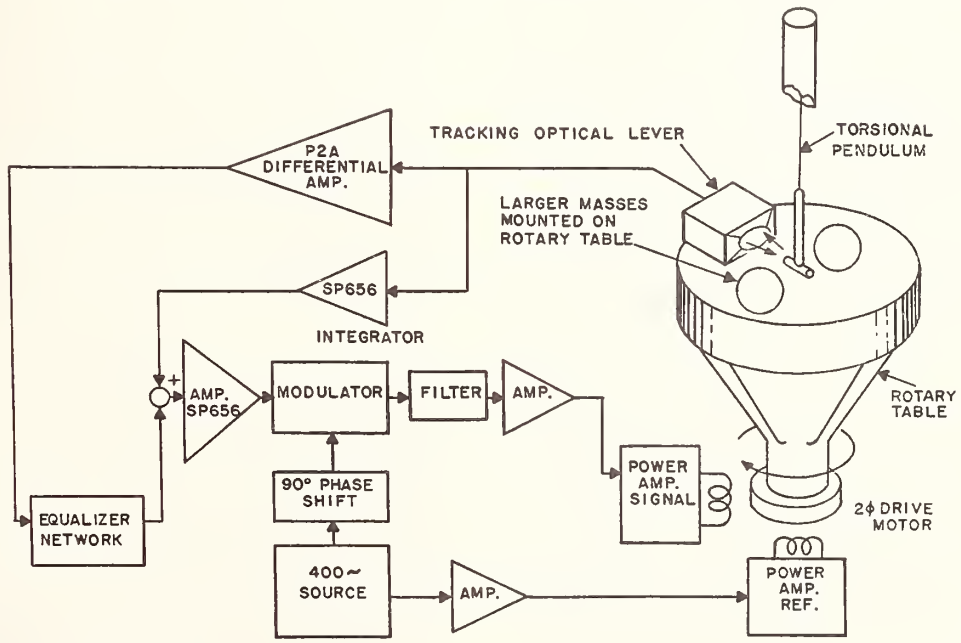


FIGURE 4. *Circuit block diagram for tracking loop.*

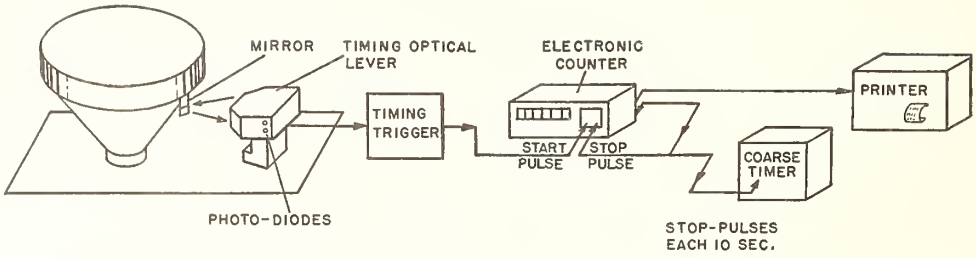


FIGURE 5. *Block diagram of timing system for measuring the period of each revolution.*

This activates the trigger when the output voltage exceeds a pre-set level. The timing trigger starts the counter (Hewlett-Packard Model 5245L) which measures the time interval from the start of the pulse until it receives a stop pulse from the counter's reference precision oscillator. The duration of this interval is relayed to a printer which records the interval in microseconds. Each 10 second interval thereafter from the reference oscillator is recorded. When the table makes a complete revolution another start pulse is produced by the same fixed optical lever which marks the beginning of a second revolution. This start pulse together with the next stop pulse from the reference oscillator will cause the electronic counter to print out the time interval between the two pulses in microseconds. The rotational period of the table is then obtained in microseeonds. The reference oscillator in the counter has been checked against a 10 MHz broadcast signal from WWV and found to be 1 MHz to within 3 parts in 10^9. The error in this trigger system even at the maximum speed of rotation of the table is less than 1 μs. This introduces an error of less than 10^{-7} radians or a precision of at least 1 part in 10^7 for the period of rotation of the table. After the successive periods of rotation of the table are known it is possible to compute the constant acceleration α of the rotary table.

The procedure in carrying out this experiment consists in first allowing the apparatus in the room to come to temperature equilibrium with the servo system operating and the rotary table stationary. The table is leveled and the run out checked by an electronic indicator gauge which records a difference in position of 7.6×10^{-6} cm. The table wobble is less than 0.0005 cm at a radius of 12 cm. The electronic indicator is next used to check the vertical alignment of the cylindrical chamber surrounding the small mass system. It has a maximum run out of less than 0.0006 cm. The quartz fiber is adjusted to coincide with the axis of rotation to 0.0005 cm by observing it with a microscope through ports (with the glass removed). The windows are replaced and the chamber is filled with helium gas. As previously mentioned the tracking optical lever is next set so that the effective twist in the fiber approaches as close to zero as possible. The tungsten spheres which already are at room temperature are placed in their mountings and their heights adjusted until their centers of mass are in the same horizontal plane (±0.002 cm) as the axis of the cylindrical rod of the small mass system. An analysis shows that the error introduced by this misalignment is negligible in the present experiment. The centers of mass of the two tungsten spheres are placed in a line through and perpendicular to the axis of rotation by use of a combination of an interferometer plate mounted on top of the chamber, modified optical lever and electronic indicator. The electronic indicator is then used to make the radial distances of the two spheres from the axis of rotation equal to within $\pm10^{-5}$ cm. The absolute value of R is determined by comparison with Johansson gauges by standard optical methods used in connection with the electronic indicator.

The tracking is then started and the acceleration of the table determined. When operating properly this angular acceleration of the table is constant. In fact, the constancy of this acceleration is a good measure of the noise in the system. After the acceleration (between 4 and 5×10^{-6} rad/s^2) takes place for several hours and the rotational speed of the table reaches between one and two revolutions per minute, the large tungsten spheres are carefully removed from the table and the tracking continued for several hours. The acceleration of the table with the spheres removed $\dot{\omega}_{w/o}$ gives a measure of the gravitational interaction of the small mass system with the mass of the table and other fixed masses turning with it as well as of fiber twist and other perturbing factors. Consequently, the acceleration $\dot{\omega}$ due to the tungsten spheres or large mass system alone is determined by eq (20). Experiments have been carried out with the table decelerating as well as accelerating. It is important to note that in all of the experiments the table rotates a number of complete revolutions; this renders negligible the effect of the ever present fixed mass asymmetry surrounding the experiment.

5. Results

Preliminary values obtained for G some months ago [4] revealed several sources of uncertainties which were limiting the precision of the measurements. Consequently, major emphasis has been placed on improving the various parts of the experiment and developing better procedures for performing delicate measurement tasks.

In the earlier experiments, the major error occurred in determining $\dot{\omega}$ and resulted from a combination of variations in spindle friction, some instabilities in the electronic circuits, and temperature gradients pro-

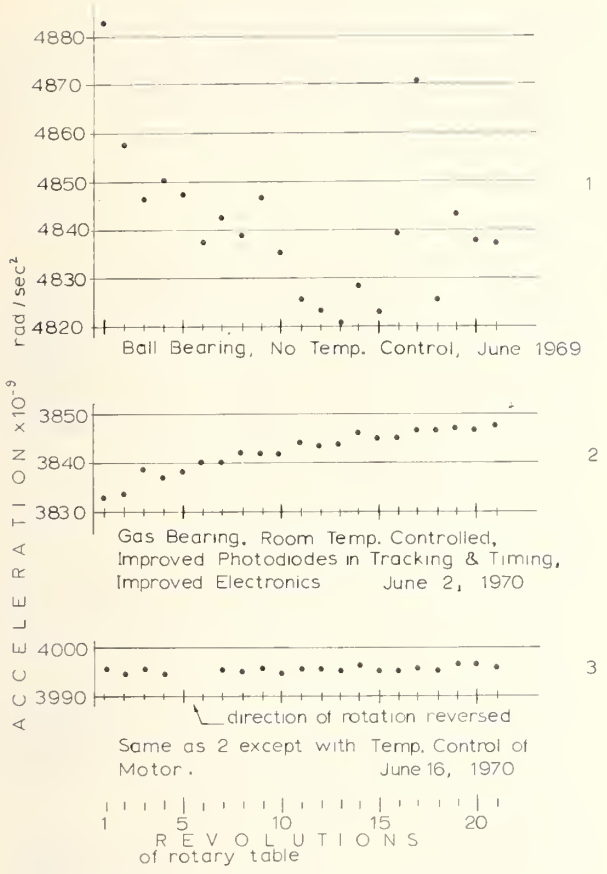


FIGURE 6. *Measurements of angular acceleration* α.

duced by the motor and bearing. Fortunately the trouble with the bearing has been solved by substituting a gas bearing for the precision ball bearing spindle used previously. Careful thermal isolation of the motor has reduced temperature variations and the electronic circuits have been much improved and stabilized. Very recent measurements of acceleration are shown in figure 6 and compared with earlier results. The acceleration now remains constant to the order of one part in 10^4. There is good reason to believe that this can be improved by at least a factor of ten in the near future. When the measurement of all the required physical dimensions of the system are completed a significant improvement in the accuracy of G should result. Before the recent improvements, our best value obtained with the present technique was $G=(6.674\pm0.012)\times10^{-11}$ Nm^2/kg^2 where 0.012 represents three standard deviations. This agrees very well with the value due to Heyl [9] of $G=(6.670\pm0.015)\times10^{-11}$ Nm^2/kg^2.

One limiting factor to the precision with which G can be determined with the present equipment is the precision with which the center of mass of the two tungsten spheres can be determined. These uncertainties produce an uncertainty in R which will produce an uncertainty in G of about 5 parts in 10^5. We are encouraged to believe that the method can be made to measure G with much greater accuracy than is possible with the present tungsten spheres.

Experiments [10] have shown that certain metals with intermediate to high specific gravities may have variations in density of less than one part in 10^7 per cm. These metals cannot be made into precise spheres because they will distort on the necessary supports under their own weight. However, it can be shown that they can be used in the form of rectangular blocks where the distortion will be greatly reduced. This of course complicates the theory but the problem is solvable by a computer.

It is difficult to estimate with reliability the ultimate precision with which G can be measured by this method. In agreement with the original conception of the experiment [3], it is still believed that magnetic suspension of the small mass system, rather than by quartz fiber, will provide the better ultimate accuracy. The primary reason is that almost surely the torsion constant of the small mass suspension system can be very significantly reduced and thus the required tracking tightness can be reduced. Also, it can be expected that the temperature sensitivity of the torsion constant can be reduced. Our experience has shown that many important improvements can be made in the present apparatus. With a completely new design of the rotary table and chamber with special attention to temperature and vibration isolation, certain changes in the servo motor drive system, improved metrology, substitution of the blocks of metal for the sintered tungsten spheres as well as of an accurately ground quartz cylinder in the small mass system, etc., it is not beyond conception that G can be measured to one part in 10^6. Also there is better reason to believe that variations in G of one part in 10^6 ultimately may be detected by this method.

6. Acknowledgments

We are especially indebted to Dr. James Senter and Dr. R. D. Rose for valuable contributions to the initial experiments. Also we wish to thank Dr. A. G. McNish and the members of the Metrology Section of the National Bureau of Standards for helpful advice, as well as Mr. R. Hibbs and Mr. E. W. Bailey of the Union Carbide Nuclear Company for assistance in obtaining the tungsten spheres. This work has been supported in part by a grant from the National Aeronautic and Space Administration.

7. References

[1] Cook, A. H., these Proceedings.
[2] Champion, F. C., and Davy, N., Properties of Matter (Blackie and Son Ltd. Glasgow, Scotland 1959) 3rd ed.
[3] Beams, J. W., Kuhlthau, A. R., Lowry, R. A., and Parker, H. M., Bull. Am. Phys. Soc. **10**, 249 (1965).
[4] Rose, R. D., Parker, H. M., Lowry, R. A., Kuhlthau, A. R., and Beams, J. W., Phys. Rev. Letters **23**, 655 (1969).
[5] Rose, R. D., A New Method for Determination of Newton's Gravitational Constant, Dissertation, University of Virginia, August 1969.
[6] Jones, R. V., and Richards, J. C. S., J. Sci. Instr. **36**, 90 (1959).

[7] Nash, J. H., Neely, A. C., and Steger, P. J., A.E.C. Research and Development Report No. Y–1654, Oak Ridge Y-12 Plant, Union Carbide Nuclear Co.
[8] Towler, W. R., and McVey, E. V., Control System for Use in the Measurement of Ultrasmall Torques, IEEE Transactions on Automatic Control (April 1969).
[9] Recommended Unit Prefixes; Defined Values and Conversion Factors; General Physical Constants, National Bureau of Standards Miscellaneous Publication No. 253 (U.S. Government Printing Office, Washington, D.C., 1963).
[10] Beams, J. W., Bull. Am. Phys. Soc. Series II, **11**, 526 (1966).

DISCUSSION

E. R. COHEN: I have some questions about how you really get one part in a million out of all of this. With the present system, I looked at your paper and tried to make some estimates. Going to the assembled masses I think would certainly help identify the center of mass properly. But in terms of the small mass, the attracted cylinder, don't you also have to worry about the density distribution in that? And I'm also worried about how many terms in your harmonic expansion are really needed to get a part in a million. I estimated, I think, at one time keeping up to the eighth order terms.

H. M. PARKER: I have done up to tenth order. I don't think that's the problem. After all, that's just work. You can do that.

E. R. COHEN: Yes. But there is the question of the interaction, of how accurately you know one of the lower order terms compared to the higher order term that you're keeping.

H. M. PARKER: Let me point out to you, though, that there is a great tendency for the mass in the small system to cancel out in the moment of inertia term and in the interaction term. If you consider your small mass system as a simple point dumbbell, then the mass doesn't appear at all. It's just the physical dimensions of it that count. I personally think that the thing that really counts here is the length of that cylinder and then—

E. R. COHEN: It's not the mass as much as the distribution, the density distribution.

H. M. PARKER: Well, I think that one can show that even in that case that same tendency is there to some extent at least. I haven't shown that yet, but I think it's there. But I think that one can build systems like this, small things like this, with densities that are uniform to a part in 10^6, don't you?

E. R. COHEN: I'm asking you. (*Laughter*).

H. M. PARKER: I'm not a metallurgist.

E. R. COHEN: I agree it's not the moment of inertia and it's not the mass. It's the radius of gyration and the higher moments of this quantity.

H. M. PARKER: Yes. I think one of the nastiest things here, if I can just correct for the moment of inertia and the interaction too eventually . . .

E. R. COHEN: Yes.

H. M. PARKER: It would be real nice if that were completely axially symmetric to eliminate one of them.

D. HALFORD: Is the distance of the large masses self-calibratable by translating each one in the X, Y, Z directions and letting it decide its own distance?

H. M. PARKER. I don't know. It could very well be. I'll tell you what we worried about. Incidentally, it's not R that really counts. It's $2R$. It's the distance between those two masses if you assume they are approximately symmetrically placed.

D. HALFORD: Yes. Each one would have to be done individually.

H. M. PARKER: Yes. And the way that we hope eventually to get at something about this non-uniformity of these tungsten spheres is to change their orientation. So far we have been careful to put them on the table the same way every time. But eventually we'll start turning them to different orientations and see if we get effectively different results. That will give us something about the inhomogeneity. There is likely to be fairly—I'm afraid too—large inhomogeneities in these sintered tungsten spheres. They were hard to mak..

PANEL DISCUSSION:
SHOULD LEAST-SQUARES ADJUSTMENTS OF THE FUNDAMENTAL CONSTANTS BE ABOLISHED?

Handling of Discrepant Data in Evaluations of the Fundamental Constants

P. L. Bender*

Joint Institute for Laboratory Astrophysics of the University of Colorado and the National Bureau of Standards, Boulder, Colo. 80302

Adjusters of the fundamental constants often face situations where different measurements of the same quantity are discrepant. Such a case may be handled by rejecting some of the measurements or by expanding the errors to the extent necessary to achieve consistency. Each of these procedures has serious disadvantages. A compromise procedure is proposed in which the evaluator estimates probabilities for the correctness of each of the possible methods of handling the data and arrives at a final mean and standard deviation using appropriately weighted results of all the methods.

Key words: Discrepant data; fundamental constants; least squares adjustments; standard deviation.

When the uncertainties assigned by the experimenters seem reasonable on physical grounds but the results clearly disagree, one point of view is that there is no meaningful way in which we can assign a most probable value and an uncertainty. There is a considerable amount of truth to this statement, but I don't feel that it should lead us to try to stop the publication of evaluations. Most users of the constants are not in a position to go through all of the relevant material themselves. They would rather use values which are at least consistent with each other

* Staff member of Laboratory Astrophysics Division, National Bureau of Standards, Boulder, Colo. 80302.

and which someone who has studied the problem carefully feels are as close as one can come to the true ones at the time of their adoption. The majority of users want a single value and an uncertainty to use in their work, and anything more complicated will lead them to go to someone else's evaluation.

The major question is how to make the value and uncertainty that are chosen as unbiased as possible. Up until now people doing evaluations have felt compelled to make crisp yes-or-no decisions on how to handle discrepancies, and no allowance has been made in the final uncertainties for the possibility that the wrong choice was made. I think that the time has now come when we would do better to recognize the degree of subjectivity and haziness which actually exists and to give estimates of the probabilities for different ways of handling the data being correct. Deciding on the final value and uncertainty then becomes easy.

Let's take an example. Consider results which fall into two groups, with perhaps four values in the higher set and two values in the lower one. Either the higher or the lower set alone would give a small chi-squared value, but together they give perhaps three times too high a value. There are no other grounds for distrusting any of the results. The three hypotheses which suggest themselves are: (A) the high values have systematic errors and should be dropped, (B) the low values have systematic errors and should be dropped, and (C) the possible errors in most of the experiments have been underestimated and the results happen by chance to fall in two groups. In this last case, one would want to increase the error bars on all the results until chi-squared was reduced to a value which would be exceeded perhaps 70 percent of the time.

I don't see any reason for the evaluator to pick one of the above hypotheses as the only possible one and ignore the others. Instead, I would have him estimate probabilities for each of the hypotheses being correct, and state what his estimates are. Under each hypothesis the probability distribution for the actual value would be taken as a Gaussian with the proper mean and standard deviation. These distributions would be weighted by the estimated probabilities for the hypotheses, and the resulting overall distribution would be used to calculate a final mean and standard deviation.

This procedure certainly does not remove the effect of subjectivity, but I believe that the results would be as dependable and unbiased as the particular person doing the adjustment is capable of. I also suspect that the amount of difference between evaluations done by different people would be reduced. One person might weight the hypotheses 60 percent, 20 percent, 20 percent, while another might prefer 40 percent, 20 percent, 40 percent. However, this seems better than requiring 100 percent, 0 percent, 0 percent type choices. For example, if the low values were thrown out, and no allowance made for this choice being wrong, it is clear that the actual value would be: (a) more than 50 percent likely to be below the quoted value, and (b) more than 30 percent likely to lie over one standard deviation from the quoted value. With the suggested procedure these biases should be removed.

One technical point may be worth mentioning. Within the constraint of picking one number for the recommended value and one for the error measure, one could also argue for picking a 70 percent confidence interval such that there is 15 percent probability of the actual value being below the interval and 15 percent probability of it being above the interval. The recommended value would then be the midpoint of the interval. This method seems somewhat less desirable to me because the recommended value would be biased. However, either method would give a reasonably fair estimate of the actual uncertainty in the result, which is the most important goal.

To summarize, the basic desire is to find a way to avoid having the quoted uncertainty in the results be systematically too small because of throwing out data, and to avoid having the recommended value be biased in a known direction. The specific suggested procedure is:

1. Formulate two or three hypotheses on how to handle the data.
2. Estimate the probability of each hypothesis being correct.
3. Calculate the resulting mean and standard deviation.

Any procedure which accomplishes the above goals would seem worthwhile.

A set of constants is needed, 'tis true,
And we must give adjusters their due.
But why can't they see
How better it would be
If they'd used my experiment too!

B. N. Taylor

Comments on Least-Squares Adjustments of the Constants*

B. N. Taylor

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Some of the author's personal opinions, as formed during the preparation of the 1969 constants adjustment based on the Josephson effect measurement of $2e/h$, are given with regard to the general utility of carrying out adjustment studies, methods for handling inconsistent or discrepant input data, how errors should be treated and assigned, how experiments should be reported, etc.

Key words: Discrepant data; fundamental constants; least squares adjustments.

1. Introduction

Ever since R. T. Birge first pioneered least-squares adjustment studies of the constants in the late 1920's, the field has been somewhat controversial. No doubt part of this has arisen from the critical comments constants adjusters must necessarily make about various experiments. However, some of the criticism has centered around valid problems such as the general utility of carrying out adjustment studies, the best methods for handling input data, particularly inconsistent data, how errors should be treated and assigned, etc. It is the purpose of this paper to summarize briefly some of the author's personal opinions on these subjects as formed during the preparation of the 1969 adjustment based on the Josephson effect measurement of $2e/h$ [1].

2. Should Least-Squares Adjustments Be Abolished?

The first question to be considered is whether least squares adjustment studies of the constants serve a useful purpose. My answer to this question is an emphatic yes, but perhaps not for the most obvious reasons. True, it is worthwhile to have at any given epoch a consistent set of constants which can be used by all workers requiring them. But this is the *least* important result of a constants adjustment—the most valuable contribution of such studies to human knowledge is the information gained during the course of the critical review which necessarily accompanies the adjustment. If properly done, this review will (a) often uncover errors in experiments and theoretical calculations; (b) reassign error estimates so that they are expressed on as equal a footing as possible; (c) find discrepancies among various measurements; (d) point up inconsistencies and weak spots in certain areas, thereby stimulating new experimental and theoretical work; and (e) summarize in one place a large amount of information which would ordinarily be available only after many hours of literature searching. That all of this is not just wishful thinking may be shown by some actual examples from the 1969 adjustment:

(1) Errors in the calculation of the Rydberg constant for infinite mass, R_∞, were discovered and the poor quality of the available experimental data emphasized. As a result, at least one new experiment to remeasure R_∞ has been undertaken.

(2) An error in the calculation of the Faraday constant from the measured value of the electrochemical equivalent of silver was uncovered.

(3) Several previously ignored experimental problems associated with what was thought to be one of the most reliable measurements of μ_p/μ_n were brought to light [2]. This has stimulated workers in the field to continue and improve their own measurements of this quantity.

(4) A mathematical error was uncovered in the theoretical calculation of the anomalous magnetic moment of the muon.

(5) The discrepancies between the theoretical and experimental values of the Lamb shift in hydrogen and deuterium were clearly demonstrated by assigning errors in a consistent way. The probable reason

* This paper was prepared while the author was with RCA Laboratories, Princeton, N.J. 08540.

for the disagreement and associated challenge to quantum electrodynamic theory has just recently been given by Applequist and Brodsky [3] as a miscalculated term in the theoretical equation for the Lamb shift.

A second and equally important reason for carrying out least squares adjustment studies of the constants is that it is one of the few ways in which the overall consistency of physical theory can be systematically investigated. The fundamental constants are important links in the chain of theory which binds all of the diverse branches of physics together, and the careful study of their numerical values as obtained from various experiments in the different fields of physics can give significant information about the overall consistency and correctness of the basic theories of physics themselves. A good example of this is the 1969 adjustment in which a low-temperature solid-state physics experiment gave information about the excited states or "polarizability" of the proton, thereby removing a serious challenge to quantum electrodynamics.

3. Importance of Numerical Values of the Constants

In light of the discussion of section 2, it may be said that the particular numerical values obtained for a final "best" or "recommended" set of constants is of only secondary importance. The attitude of the vast majority of scientists towards the constants also supports this thesis and is aptly demonstrated by a question put to me by a colleague during the course of the work on the 1969 adjustment. He asked: "Is $\hbar$ still 10^{-27}?" (which it is to within about 5%). Although obviously said in jest, this question points up the fact that most workers do not really require the last few decimal places in the values for the constants. As evidence, I cite the fact that no cry of anguish was heard from the general scientific community when it was discovered in 1967, as a result of the Josephson effect measurement of $2e/h$, that many of the previously accepted values of the constants would have to be changed by between 20 and 100 ppm (4 to 5 times their assigned errors) [4]. Indeed, there was not even as much as a whimper! Handbooks continued to publish the old values without commenting on their obsolescence (even as late as the 1969–1970 edition of a well-known reference book) and scientists continued to use them. To my knowledge, there have been no serious consequences. (In browsing through the 1956–1957 edition of the same book, it was discovered that the 1941 recommended constants of R. T. Birge were included along with the more modern values, presumably so that the reader would have a choice!) The fact to keep in mind is that those scientists who really need to use the last decimal places will not be content just to take numbers out of a table but will go to the originating article. Those workers who are content to use the numbers as given without worrying about where they came from could use almost any number.

4. Treatment of Discrepant Data

In light of the discussion of section 3, the problem of what to do with discrepant or inconsistent data takes on less importance. (By discrepant data is meant, for example, two measurements which differ from each other by between two and three (or more) times the standard deviation of their difference.) Indeed, the actual method to be used to handle such data is essentially a matter of personal taste. The one requirement is that what has been done be clearly described and the results of alternate approaches be summarized.

Some actual methods for treating discrepant data in a least squares adjustment might include:

(a) Use as input data all apparently reliable measurements even though they may be inconsistent, but expand errors sufficiently so that a satisfactory value of χ^2 results. (In practice, depending on the situation, the errors of the subset of inconsistent data may be expanded, or the errors of some or all of the input data may be expanded.)

(b) Expurgate the inconsistent measurements thought to be least reliable and somehow expand the errors of the affected adjusted output values to take into account the expurgated data.

(c) Decide on as sound an experimental and theoretical basis as is possible which of the inconsistent data is least reliable and expurgate them, but expand no errors.

For two main reasons, I believe this last solution is generally the best. First, because expanding errors of quantities is always somewhat arbitrary and subjective, it is difficult to justify objectively. In a way, it is an admission of defeat—that no experimental or theoretical basis for making a rational decision can be found and that the only way out is to assume that all the data is bad. It therefore throws away information by making quantities more unknown than they may actually be. Second, it seems to me that people will be more apt to make new calculations or to repeat a measurement (or devise new ways to do an experiment) if the discrepancies are dramatized by censoring out apparently inconsistent data. The challenge to prove the constants adjuster incorrect, thereby forcing a new adjustment, is much more stimulating than the challenge to obtain a number which will just be thrown into the "pot" and which will probably have little effect (as would be the case if all the inconsistent data had been included with expanded errors).

An example of the usefulness of this solution is contained in the 1965 adjustment of Cohen and DuMond [5]. The main problem in this adjustment was the disagreement between two different values of the fine structure constant. The deuterium fine structure measurements of Lamb and co-workers gave $\alpha_D^{-1} = 137.0388(6)$ (4.4 ppm) and the hydrogen hyperfine splitting measurements of Ramsey and co-workers gave $\alpha_{Hhfs}^{-1} = 137.0352(17)$ (12 ppm), where the error in the latter value was liberally

assigned to take into account uncertainties in the theoretical equation for the Hhfs due to proton polarizability. Cohen and DuMond chose to discard $\alpha_{\mathrm{Hhfs}}^{-1}$ because values of α^{-1} derived from the muonium hyperfine splitting and anomalous moment of the electron, although of larger uncertainty, agreed with α_{D}^{-1}.

As a result of Cohen and DuMond's decision, subsequent tests of quantum electrodynamic theory using the 1965 recommended adjusted value of the fine structure constant (i.e., α_{D}^{-1}) were more incisive than if the recommended value had been representative of the mean of α_{D}^{-1} and $\alpha_{\mathrm{Hhfs}}^{-1}$, with an appropriately expanded error. That is, by rejecting $\alpha_{\mathrm{Hhfs}}^{-1}$, they forced the separate use of both values in any critical application. The dilemma was thus brought to the attention of the scientific community, stimulating much new theoretical work aimed at a better understanding of the hyperfine splitting in hydrogen. (This work included important contributions by Iddings [6], Brodsky and Erickson [7], and Drell and Sullivan [8].) Although the Josephson effect measurement of $2e/h$ subsequently showed that $\alpha_{\mathrm{Hhfs}}^{-1}$ was the more nearly correct value, it is now clear that Cohen and DuMond's incorrect decision served a very useful purpose. Indeed, I would claim that it served a *more* useful purpose than would have the decision to use some kind of average value with an expanded error.

It is of course obvious that the price one pays for the solution I advocate is that if it later turns out that the wrong data have been expurgated, then the adjusted values of certain constants may change by one or two standard deviations or possibly more. (A glance at the first and last figures in reference [1] will show that this has been the rule rather than the exception.) Many people find this possibility disturbing (e.g., see the paper by Bender, these Proceedings), and favor some type of procedure which includes all of the data so that the final values will be representative of all the data, hopefully resulting in more accurate final values (i.e., less subject to future change). But there is no guarantee that these values will in fact be more reliable! On the other hand, it is probably true that such constants can be used uncritically with more safety. But this attitude is analogous to requiring that all knives be sold with blunt edges so that users who do not wish to take the trouble to learn the proper use of a knife can nevertheless use one without hurting themselves. Our philosophy in the 1969 adjustment was to provide the best possible cutting tool for those workers who needed it most and who were prepared to use it in full knowledge that care was necessary to avoid cutting themselves. In short, since the majority do not particularly care what numbers they use, adjustments should be geared to the small but more important minority who do and who need the most useful and stimulating numbers they can get their hands on.

But I should hasten to add that no matter what method is used to handle discrepant data, in order to bring home to the average worker that a set of recommended constants is not inviolate and handed down on stone tablets, every table of constants, whether original or reprinted, should probably start off with some type of warning label (preferably in large bright red letters) such as:

Warning!

Because of the intimate relationships which exist among least-squares adjusted values of the fundamental constants, a significant shift in the numerical value of one will generally cause significant shifts in others. Consequently, for any critical application of these numbers, the user is urged to refer to the originating article as well as the current literature in order to ensure that he understands their limitations and whether they are not in fact totally obsolete.

5. Reporting Results

Because of the importance of clearly reporting experimental and theoretical results, I would like to conclude with a restatement (essentially verbatim) of certain ideas and recommendations given in [1].

During the preparation of the 1969 adjustment, it became apparent that our task would have been much easier if certain procedures had been generally followed in the reporting of experimental and theoretical results. More important, we would have been able to compare theory with experiment and to incorporate data into our least-squares adjustment with a great deal more confidence. We therefore have several suggestions for presenting experimental results which, if followed, would facilitate the work of future reviewers and materially increase the reliability of direct comparisons between theory and experiment and the output values of any least-squares adjustment.

(1) It should be made quite clear exactly what steps were taken to arrive at the final result from the actual experimental observations. If at all possible, the data analysis should be presented in such a way that the reader can follow each important step of the analysis and duplicate the final reported results. This means that particular attention must be paid to the way in which presumably well-known constants and corrections of various kinds enter the calculations. The experimenter should ask himself, "*Have I provided enough information in a sufficiently clear manner so that 10 years hence my result can be updated in light of any new information or data which may become available?*"

(2) Every effort should be made to discuss in some detail each important systematic error believed to be present in the experiment and to list estimates of these errors in tabular form, preferably as 70 percent confidence-level estimates (i.e., standard deviations). We strongly deplore statements such as "the systematic error is not believed to exceed 0.5" without any further discussion of the individual systematic errors contributing to this value or whether the quoted value is meant to be a 50 percent, 70 percent,

or 95 percent confidence-level estimate. We also urge that the statistical uncertainty be computed using standard methods and presented in the form of a standard deviation. Since the uncertainty assigned a particular experimental datum plays a major role in determining whether it is in agreement or disagreement with theory and how much weight it will carry in a least-squares adjustment, correct error estimates are of the utmost importance.

(3) The numerical value of a datum and its uncertainty should be presented with a sufficient number of significant figures to prevent rounding errors when it is used. As a rule of thumb, the error should usually be stated to two-digit accuracy, and the datum itself should contain enough digits so that the quoted uncertainty corresponds to the last two figures.

(4) When the experiment is completed and the results published, the original laboratory notebooks and other pertinent information should be deposited in duplicate in secure places; use of a departmental safe or bank vault is not presumptuous, but prudent. No matter how detailed the published report may be, these items constitute the only really complete record of the results of the expenditure of man-years of effort and many thousands of dollars of someone's money. Their preservation may prevent much uncertainty and unnecessary future expenditures.

Lest our theoretical colleagues feel slighted, may we suggest that when they write up the results of their calculations, they do the following: (1) Clearly summarize the final results in one place so that the reader doesn't have to search through innumerable pages. (2) Discuss the effects of the important approximations and/or those of uncalculated terms. (3) Present the results in such a way that corrections can easily be made at a later date for changes in any of the numerical constants used in the calculations. (4) Discuss any differences, even if they are apparently insignificant, between their results and those of other workers. (5) Keep in mind that experimentalists may wish to read and understand the paper and use its results.

We recognize that these recommendations will be considered a nuisance by some workers, but it seems to us that anyone who can spend two or three (or perhaps five) years measuring or calculating a particular quantity can certainly spend an extra month or two writing up his results in a manner which will maximize their future usefulness. We also believe rather strongly that the responsibilities of a scientist include assuring complete and open access to his results and procedures for a considerable time after the completion of his work.

6. Acknowledgments

I should like to thank D. N. Langenberg and W. H. Parker as well as the members of the NAS/NRC Advisory Committee on Fundamental Constants for stimulating discussions.

7. References

[1] Taylor, B. N., Parker, W. H., and Langenberg, D. N., Rev. Mod. Phys. **41**, 375 (1969). Also published as a monograph under the title The Fundamental Constants and Quantum Electrodynamics (Academic Press, Inc., 1969).
[2] Fystrom, D. O., Petley, B. W., and Taylor, B. N., these Proceedings.
[3] Applequist, T., and Brodsky, S. J., Phys. Rev. Letters **24**, 562 (1970).
[4] Parker, W. H., Taylor, B. N., and Langenberg, D. N., Phys. Rev. Letters **18**, 287 (1967).
[5] Cohen, E. R., and DuMond, J. W. M., Rev. Mod. Phys. **37**, 537 (1969).
[6] Iddings, C. K., Phys. Rev. **138**, B446 (1965).
[7] Brodsky, S. J., and Erickson, G. W., Phys. Rev. **148**, **26**, (1966).
[8] Drell, S. D., and Sullivan, J. D., Phys. Rev. **154**, 1477 (1967).

The values of the physical constants
Are fixed, no matter how one rants.
But the numbers to use
Can destroy or confuse
The decorum of many a conference

E. Richard Cohen

In Defense of Least-Squares Adjustments

E. Richard Cohen

Science Center, North American Rockwell Corporation, Thousand Oaks, California 91360

In the approximation of linear functions and assuming that the best estimate of a quantity based on several observations of equal accuracy is the mean of the observations, it is shown that the usual relationships of the "method of least squares" can be derived without any reference to quadratic forms or the minimization of squares.

Key words: Consistency and invariance of constants; fundamental constants; method of least-squares; variance.

1. Introduction

I do not personally believe that the method of least-squares needs a strong defense; however, it may well profit from an examination and explanation. I do not want to emphasize the arguments for the method of least-squares but rather look at the structure of the resultant expressions. I shall concentrate on the *linearity* of these expressions and not on the *quadratic* form from which the expressions may be derived, and even less on the fact that this quadratic form has been associated with a multidimensional Gaussian probability distribution function. By avoiding such an identification it may be possible to avoid most of the inherent objections which have repeatedly been raised with regard to the justification of the use of least-squares.

To begin, let it be granted that the analysis of the numerical values of the fundamental constants has two separate and distinct objectives; that which we call "the method of least-squares" may be a useful technique for one of these and may not be a useful technique for the other. The two objectives are: (a) the analysis and critical evaluation of the available experimental data in order to determine the consistency of the data and the adequacy of the current status of theory—either the "fundamental" theory of the relationships between observables or the more pedestrian theory of the implementation of an experiment (this is the analysis that allows one to either avoid or remove "systematic" error); (b) the development of a set of numerical data which can be used as reasonably "standard" values with the assurance that they are consistent among themselves and with the consensus of experimental information upon which they are based.

In objective (a) there may be several statistical procedures which could and should be used, but there are conditions associated with objective (b) which almost uniquely prescribe the method of least squares.

The appropriate procedures for treating numerical data will depend on the end use to which that data is to be put. If one knows precisely what that end use is to be, one can often optimize the procedures to fit that purpose. In the case of a recommended list of general physical constants, the end uses are broad and the controlling factors are consistency and invariance. By consistency I mean the numerical agreement with all known theoretical relationships. A prime example of this is the requirement that the numerical values of m, e, h, and c, although individually known to much lower accuracy, should when put into the Bohr formula give the correct value of the Rydberg constant as well as the correct (essentially zero) error, since this relationship is an imposed auxiliary condition on the data. By invariance I imply that the numerical values of the constants should be independent of the choice of those quantities which were chosen as the primary unknown variables of the adjustment.

The ability to achieve these requirements is greatly simplified in the present adjustments because the errors are small enough (of the order of parts per million) that one need only introduce first derivatives in the Taylor expansions of any functions of the

variables in the evaluation of errors. Thus the theory of the adjustment can be reduced to a completely linear theory and one can consider all functions to be linear transformations which we shall write for example as $f_i = A_{i\alpha} x_\alpha$ (summation understood on repeated indices) or simply as $f = Ax$.

2. The Overdetermined System of Equations

The primary question is how to treat an overdetermined set of equations [1],

$$y_\nu = C_{\nu\alpha} x_\alpha + \epsilon_\nu, \tag{1}$$

where y_ν represents the experimental observations, the matrix C is a set of well defined coefficients, x_α the variables to be fit, and ϵ_ν is the unknown error in the observed numerical value. Let σ_ν be the estimated or assigned standard deviation of the error. We want to obtain an adjusted set of values $\hat{x}_\alpha$ and an estimate of the variance, $\mathrm{var}(x_\alpha, x_\beta)$, which are unbiased and consistent. The numerical results for our adjustment should not depend on whether we use as variables e, h, and N or, for example, α, F, and μ_p/μ_n. In particular this requires that, had we used a different set of variables, $f = Ax$, we could calculate $\hat{x}_\alpha$ and $\mathrm{var}(x_\alpha, x_\beta)$ equally well from

$$\hat{x} = A^{-1}\hat{f} \tag{2}$$

$$\mathrm{var}(x_\alpha, x_\beta) = A^{-1}\,\mathrm{var}(f_i, f_j)\,A^{-1^T} \tag{3}$$

Note that the expression for the variance, which is the so-called "generalized equation for the propagation of error" does not in itself depend on any assumption of least squares; it is entirely a result of general properties of error distribution functions and requires only that the distributions have finite second moments.

Because of eq (2) we are restricted to using a *linear* expression for $\hat{x}_\alpha$ in terms of the experimental data

$$\hat{x}_\alpha = B_{\alpha\nu} y_\nu \tag{4}$$

where $B_{\alpha\nu}$ is a matrix to be determined. In general, since the experimental data forms an inconsistent overdetermined system, the matrices B and C are nonsquare and are not inverses. Although $BC = I$, there is no similar relationship for CB. In fact the estimates of the errors ϵ_ν are given by

$$\epsilon = (I - CB)y. \tag{5}$$

In the simplest example, repeated observations of equal accuracy of a single variable, one has

$$y_\nu = x + \epsilon_\nu. \tag{6}$$

Since there can be no preference to the ordering of the data set, $\hat{x}$ must be a symmetric function of y_ν and since it must also be linear, we will have

$$\hat{x} = (1/n) \sum y_\nu. \tag{7}$$

If we ascribe weights to each of the observations, this becomes

$$\hat{x} = \sum w_\nu y_\nu / \sum w_\nu. \tag{8}$$

In eq (1) however the various y_ν may have different physical dimensions; it is therefore in general meaningless to talk about equal accuracy, and one must use a weighted average of the observations. This is achieved by multiplying eq (1) by a matrix D^T. The matrix must be chosen in such a way that a solution is possible, i.e., such that $D^T C$ has an inverse. Then the matrix B in eq (4) will be

$$B = (D^T C)^{-1} D^T. \tag{9}$$

The matrix D^T is not unique.

The observational equations can however be reduced to a form in which all equations have equal accuracy if each equation is divided by its assigned accuracy measure. The accuracy of each equation so transformed is just unity. In such a case, the matrix D^T can be a function only of the matrix C and the choice $D^T = C^T$ is obviously the simplest possible form.

We call $D^T C$ the weight matrix of the set x_α, since it plays the same role in eq (4) as $\sum w_\nu$ does in eq (8). We now write

$$W_x = D^T C \tag{10}$$

which becomes, when expressed in terms of the unnormalized matrices

$$W_x = \sum_\nu C_{\nu\alpha} C_{\nu\beta} / \sigma_\nu^2. \tag{11}$$

Equation (11) can be immediately further generalized and put into full matrix forms by defining

$$W_y^{-1} = \mathrm{var}(y_\nu, y_\mu) \tag{12}$$

and

$$W_x = C^T W_y C. \tag{13}$$

3. Conclusions

Equation (13) is just the usual expression of the "method of least squares." However, we have minimized no squares nor have we even written down any quadratic forms. Starting with the *ansatz* that the "best" estimate of a quantity based on several observations of equal accuracy is the mean of the observations, we have generalized to the usual "least squares" result for a multivariate linear system. We have indeed been guided by the result which we wished to achieve but that result nevertheless appears to be the simplest and most direct generalization which preserves the invariance of the solution either to the basis vectors of the variable-space or to the form in which the observational data may be expressed. The result has the required invariant structure of course, only in the approximation of linear functions, but for the intended applications this is no real restriction.

The results are probably not original; if they were

not explicit in the work of Gauss himself, they are probably implicit in it. The generality of the relationship and its independence of any underlying probability distribution of errors however should be emphasized. There are left unresolved two questions: (a) whether an average is indeed the "best" estimate; and, (b) exactly how the variance matrix $\mathrm{var}(y_\nu, y_\mu)$ or weight matrix W_y, ought to be most appropriately defined. The latter question is certainly basic and represents the fundamental art in the process of adjustments. Once it is resolved however the problem is reduced to numerical calculation. The former question is also vital; although the average may not be best for some specific instance it seems to be the most useful and convenient assumption and perhaps the only one which provides the general symmetry which is required for the consistency of a generally useful "best" *set* of values.

4. Reference

[1] Cohen, E. R., Crowe, K. M., DuMond, J. W., *Fundamental Constants of Physics* (Interscience, New York, 1957) Chapter 7.

After what Branscomb has said,
Least squares "adjustments" are dead.
We'll "evaluate" now,
But continue somehow
To rush in where angels won't tread.

JOHN S. THOMSEN

Some Aspects of Least-Squares Adjustments of Constants

John S. Thomsen

Department of Physics, The Johns Hopkins University, Baltimore, Md. 21218

The need for a least-squares adjustment to obtain recommended values for fundamental constants is reviewed and its advantages briefly discussed. It is emphasized that in work of the highest precision no investigator should use the output values of such adjustments uncritically. The question of the best form of error criterion is considered; when systematic effects are predominant, probable error appears to be the most intuitive concept and the one least likely to be misinterpreted. The problem of rejection of discrepant data is discussed; it may be preferable to increase the error and thus lower the weight of such an item, rather than to reject it entirely. An example is given in terms of μ_p'/μ_n, the magnetic moment of the proton in units of the nuclear magneton.

Key words: Discrepant data; error criterion; fundamental constants; least squares adjustments.

1. Need for Least-Square Adjustments

At the risk of some repetition let us consider briefly a few aspects of statistical adjustment of fundamental constants. First of all it is essential to have a *consistent* set of values. As a trivial example, the recommended value of h/e should be consistent with the individual values of h and e. The same principle applies to the numerous interrelated constants such as the Rydberg, the fine structure constant, the Bohr radius, the radiation constants, and many others. Obviously also, the recommended set should be made as accurate as possible in the light of existing experimental data and theoretical knowledge.

So far the best procedure appears to be a painstaking critical review of all the data, followed by a least squares adjustment of the most significant items. The critical review is, of course, the most important aspect; the adjustment itself is a more routine procedure. The least squares technique or any form of statistical adjustment is occasionally criticized, but thus far no good substitute appears available. Furthermore a least squares adjustment shows clearly any experiments in serious disagreement with the rest of the data and focuses attention on these discrepancies. Hence the answer to the topic question is clear: Least-squares adjustments should *not* be abolished.

However, one should still be cautious in the use of a recommended set of constants derived by such an adjustment. Despite all of the mathematical techniques and the computer computations involved, the results ultimately depend on the validity of at least the major part of the input data, including assigned errors. Such data can become obsolete quite rapidly, in some cases almost as soon as an evaluation is in print. Furthermore adjustments are frequently based on a censored set of data with one or more discrepant items omitted from the final computation. The validity of such omissions may be open to question and the situation may change as new information becomes available. Hence in work of the highest precision no investigator should use any recommended set of values for the constants uncritically without some study of the adjustment and of any subsequent relevant data.

2. Error Criterion

Let us examine two specific related aspects of such adjustments, viz., statement of error and treatment of discrepant data. The problem of an error criterion would be rather trivial if all errors were predominantly statistical; one could then adopt standard deviation, probable error, or any other conventional measure, knowing that these differ from one another only by numerical factors. However, in many, perhaps in most cases, the major contribution to the estimated error comes from systematic effects. One can still usually form a reasonable error estimate of some sort, but the concept of a complete error distribution curve becomes rather illusory.

How then do we find a reasonable basis for assigning a numerical value to a systematic error? In most

cases, the best realistic answer seems to be that the experimenter must recognize that he is quoting betting odds. Thus, if he states a result as (100 ± 1) cm (probable error) he is asserting that there is a 50 percent probability that the true value lies between 99 cm and 101 cm. If he has formed his error estimate honestly, avoiding both overoptimism and undue conservatism, *he should be willing to take either side of the bet*. This is the essence of an honest error estimate.

The definition of standard deviation implies a knowledge of the entire error distribution. Hence, in light of the above discussion, it calls for a whole table of betting odds for various possible errors. This is a much more sophisticated requirement than probable error and seems much less realistic. Of course, it can be interpreted simply as a 70 percent confidence interval, i.e., the experimenter feels that there is a 70 percent probability that the true value lies within a distance σ of the mean. However, in this narrower sense, it offers little advantage over the more intuitive concept of probable error and may lead to unwarranted interferences. Hence the simpler concept of probable error seems preferable whenever systematic effects are important.

In either case a complication arises when there is a skew probability distribution. (The fact that the experimenter recognizes the distribution is skew need not imply that he knows it in any detail. For example, he may know the sign of a given systematic error, but be quite uncertain as to its magnitude.) In this case a result could be stated in the form

$$x=(100.0_{-0.8}{}^{+1.2})\text{ cm}\qquad\text{(P.E.)},$$

which would imply that there is a 25 percent probability that the true value lies between 99.2 and 100.0 cm and an equal probability that it lies between 100.0 and 101.2 cm. However, it is more common procedure to quote it in a form such as

$$x=(100.0\pm1.0)\text{ cm}\qquad\text{(P.E.)},$$

implying merely a 50 percent probability that the true value lies between 99.0 and 101.0 cm. This statement may be just as correct as the former. However, unless carefully qualified, it can lead to incorrectly inferring a symmetric error distribution. Hence, if the second form is used for an asymmetric error distribution, some warning statement in the text appears necessary.

3. Treatment of Discrepant Data

Let us now turn to the problem of censored data. If an experiment can be rejected for good cause *before* any adjustment is carried out, there is no serious problem. However, if it is rejected only *after* a preliminary adjustment, primarily because it disagrees with the bulk of the other data, this poses a more difficult question. Even if a plausible reason for rejection is then discovered, one wonders whether other experiments might be similarly rejected if subjected to equally critical examination.

TABLE 1. *Values of μ_p'/μ_n, the ratio of the magnetic moment of the proton (in water) in units of the nuclear magneton.*
(See B. N. Taylor, W. H. Parker, and D. N. Langenberg, The Fundamental Constants and Quantum Electrodynamics (Academic Press, New York, 1969), pp. 115, 155; Rev. Mod. Phys. **41**, 375 (1969), Tables XIV, XVII).

Low Average (Sommer et al. and Sanders et al.)

$$X_1=2.792\ 692\pm(9.9\text{ ppm})$$

High Average (Mamyrin and Frantsuzov and Boyne and Franken)

$$X_2=2.792\ 797\pm(5.9\text{ ppm})$$

Difference

$$X_2-X_1=(37.7\pm11.5)\text{ ppm}$$

Even when it is conceded after re-examination that a particular experiment is suspect, should the adjustment then be carried out as though it never existed? Does not the rejected experiment still provide some clue as to the magnitude and direction of the error in the items retained?

One alternative is to reject nothing and simply to increase all errors sufficiently to obtain a reasonably consistent picture. However, this goes to the opposite extreme and in effect precludes any critical judgement on the part of the evaluator.

A more reasonable middle course would appear to be to increase errors sufficiently to obtain a reasonably consistent picture, but to make the error adjustment primarily or entirely in the suspect item. It is sometimes asserted that this has the same practical effect as rejecting the item. However, this is not necessarily the case.

As a specific example, let us consider the measurement of the magnetic moment of the proton in terms of the nuclear magneton, i.e., μ_p'/μ_n. There are two low values in close agreement with one another, giving a mean $X_1=2.792\ 692\pm(9.9\text{ ppm})$ (standard deviation), as shown in table 1. On the other hand there are two high values in reasonable agreement with a mean $X_2=2.792\ 797\pm(5.9\text{ ppm})$. (The data quoted are taken from Taylor, Parker, and Langenberg, who state errors in terms of standard deviations. Actually, there is a fifth, intermediate value, which will be omitted in order to simplify the example.) The difference between the two averages is (38 ± 12) ppm or more than 3σ. The statistical probability of a deviation as great as this is only about 0.1 percent, which is clearly improbable.

TABLE 2. *Weighted mean values and differences obtained with various weights assigned to X_1 and X_2.*
(All quantities have been converted to ppm difference from the value of X_1; these differences are denoted as x's.)

	Case I (Unmodified)	Case II (Low Weight for X_2)	Case III (Low Weight for X_1)
x_1	0 ± 10	0 ± 10	0 ± 18
x_2	38 ± 6	38 ± 16	38 ± 6
Δx	38 ± 12	38 ± 19	38 ± 19
$\bar{x}$	28 ± 5	10 ± 8	34 ± 6

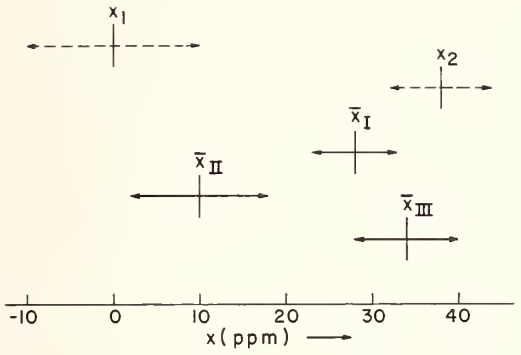


FIGURE 1. *Proton moment μ_p'/μ_n expressed as fractional deviations (ppm) from the value* x_1.

The two conflicting input values are shown as x_1 and x_2. In Case I, the mean $\bar{x}_I$ has been computed directly from the input errors ($w_i \sim 1/\sigma_i^2$). In Case II the error of x_2 has been increased to lower its weight; in Case III the error of x_1 has been increased (see table 2). The respective means are shown as $\bar{x}_{II}$ and $\bar{x}_{III}$.

Before proceeding further, let us reduce the variables to fractional deviations from X_1 in ppm (parts per million), as shown in table 2. Case I shows the results of using weights determined by the input errors ($w_i \sim 1/\sigma_i^2$). The difference between the values is (38±12)ppm, as indicated above; the weighted mean becomes (28±5)ppm, i.e., 28 ppm higher than the X_1 value.

Suppose we now say, rather arbitrarily, that a difference of 2σ is tolerable; there is a statistical probability of about 5 percent for a difference as great as this. Suppose further that the high values, i.e., x_2, are suspect. Then, as shown in Case II, we may increase the error of $\bar{x}_2$ sufficiently to reduce the discrepancy to 2σ. The mean (computed with $w_i \sim 1/\sigma_i^2$) is then in quite good agreement with x_1, but shifted toward x_2 by an amount σ_1 so that the questionable information in x_2 has not been totally rejected. The opposite situation, in which the low values, i.e., x_1, are suspect is shown as Case III; the new mean is in good agreement with x_2, but shifted toward x_1 by about $2\sigma_2/3$. The three cases are shown graphically in figure 1.

It should be noted that the errors in the various means, i.e., in the $\bar{x}$'s, are stated on the basis of internal consistency; on the basis of external consistency they should be doubled. However, in the actual adjustment of the constants those items will form only a subset of the entire data and external consistency will be determined on the basis of the χ^2 obtained for the entire adjustment; it is then unlikely to involve doubling the error.

Various modifications of the above procedure can obviously be considered. Some other criterion than a 2σ difference might well be adopted. The errors of both items might be increased, but the dubious one substantially more than the other. The main point is simply that there is an obvious middle course between uncritically raising both errors in the same ratio and rejecting one item in its entirety. Furthermore there is a significant difference in the result of the adjustment when the suspect item is downweighted, rather than totally rejected.

It is unlikely that there will ever be complete agreement in detail on the best techniques of making a statistical adjustment of the constants. However, the points discussed above seem to warrant more careful consideration than they have received to date.

Comments on the Assignments of Experimental Uncertainties

Peter Franken

Department of Physics, University of Michigan, Ann Arbor, Michigan 48105

The uncertainty assigned to a precision measurement ideally should not be solely determined by random measurement errors, but should partly be the result of subjective judgements by the experimenter on the accuracy of his work. A method based on a betting procedure is recommended.

Key words: Betting curves; experimental errors.

I am afraid I came to this conference late and I do not have a single limerick, but in listening at some of the luncheon tables I am afraid I have been reminded of a story, which I should not tell because my imitation of a Scottish accent is dreadful. At any rate, there is this Scottish military man and he—(Editor's note: In spite of its quality, this story has been deleted to protect the innocent, on the off chance there may be one among our readers.) Next time I will have a limerick.

The thing I had thought of opening with was to remind you of something a very well-known physicist (who I think attended this conference earlier) was fond of saying, namely, "There are liars, there are damn liars, and there are statisticians." Now the fact is I think he is completely wrong. I think the correct statement is "There are liars, there are damn liars, and there are just plain damn fools." There's nothing wrong with the techniques of statistics. You have seen examples of it today—the calculations are perfectly proper and the fact is that the statistical methods required for many evaluations are not really pushing the state of art of statistics.

The problem really relates—and I will get to the point that Pete has been touching on—I think the problem relates to something most of you have been exposed to either through seminars or such technical journals as *Newsweek* and *Time*, the concept of how the new information storage and retrieval systems of the next several decades are going to work. Every article that is published will be abstracted and the abstract put into a very good memory. Now a bright young man comes along and he wants to do some research. He would like to get the best idea of the current thinking on the structure of some god-awful barbarescious molecule or something like that, so he fills in some key words like "god-awful" and "barbarescious molecule." The software and hardware goes to work and a few moments later, out come some hard copies of three pertinent articles with the relevant paragraphs very nicely underlined. But the unfortunate problem is that this system will never be any better than the quality of the input abstracts. Now without meaning to hurt Lew Branscomb's feelings, it is quite clear that if this system is ever established, and these abstracts are to be prepared by an exhausted fleet of GS-9's at the Bureau of Standards, there is going to be something missing. What is needed is either a good crowd of some very sophisticated technical people to prepare the abstracts of the articles and carry out the information processing *or*, and here is my main point, the author of the article himself should be sufficiently motivated as to prepare the abstract.

Now this comes right to the heart of what I want to talk about, which is not least-squares adjustment of constants but rather the treatment of the primary data. This is rather egregiously handled by many (but, I hasten to add, by no means all) experimental physicists. What an experimentalist really should state is his estimate of a 70 percent confidence interval in which he thinks the odds are really 7 to 3. To say, "Well, I'm going to assign one statistical standard deviation," is often if not almost always absurd in atomic physics and measurements of fundamental constants. If in fact you are really limited by the random errors of your measurement you should measure some more. You want to get down to the point where you *do* have to apply subjective judgments on what you are doing.

Now the method I use is related to what Pete and others have discussed. I have used this with my own graduate students. Let us assume that the poor chap has arrived at an error by one method or another. I say to him, "Look, you've quoted, say, seven parts per million on this measurement." It is hard to visualize betting in the 7 to 3 region but it is easy to bet even odds as suggested by John. What I do is take 5/7 of whatever his figure is (in this case it would be five parts per million) and I force him to play the following game. It is the old mother-and-two-kids-home-from-school-trick to avoid quarrels—

one kid cuts the cake in half and the other kid picks his piece. I say, "Pretend that you have to wager $500—would you believe your Ph.D?—on the proposition that this is the correct error interval, and let me pick which side of the bet I take." Now this is a terribly subjective thing, but in fact it is precisely the kind of check point that should be made. There are far too many experimentalists, some of them extremely distinguished, who take the view that they are going to take two or three or n standard deviations, something large enough so that they will not be shown to be wrong. There are many experimental physicists who cannot accommodate to the idea that if they are really good, over their whole lifetime they will be wrong about 30 percent of the time. But if they can't accommodate to it then some other people, possibly less tutored in their experiment, are going to have to do the accomodation. Occasionally it might be ruthless and very often it might be incorrect. So the substance of my plea today is not really addressed to the technology of least-squares adjustments, but rather to you experimentalists out there in Disneyland. Really take a harsh look at your results using the betting concept because if you do not, you are the guys that are the damn fools!

Contribution to Panel Discussion on Adjustments of the Fundamental Constants

Churchill Eisenhart

Institute for Basic Standards, National Bureau of Standards, Washington, D. C. 20234

Continuation of least squares adjustments is endorsed. Trial of alternative adjustments determined by criteria of Least Sum of Absolute Residuals and Least Maximum Residual is suggested as checks on credibility of least squares adjustments. References are given on implementation of these alternative adjustments.

An example is given of disparate values any "averaging" of which is unwarranted. Several well known sets of data are discussed as instances in which simple "changes" revealed an important "component of variance" that needed to be taken into account both in computing "the mean" and in evaluating its reproducibility. When "systematic differences" among subsets of measurements of a single quantity are evident or revealed by statistical analysis, use of the unweighted mean of the subset means is recommended, rather than a weighted mean of these means. Bases of this recommendation are presented in detail.

It is concluded that nonoverlapping of the limits of uncertainty assigned to two or more determinations of the same quantity stems primarily from gross underestimation of the standard deviations of the respective "means" due to (a) insufficient "sampling" of relevant conditions in the experimental programs and (b) use of computational procedures that fail to take proper account of "between subsets" components of variance.

Key words: Combination of observations; components of variance; least maximum residual; least sum of absolute residuals; least sum of squared residuals; linear programing; methods of adjustment; physical constants; systematic errors; underestimation of standard deviations; unweighted mean; weighted mean.

1. Should Least-Squares Adjustments Be Abolished?

We are asked this question in the program caption of this round table discussion. My answer is an unequivocal "No."

First, from a purist viewpoint, it is essential, at least logically, to have available at successive stages of the development of physical science at least one set of values of the fundamental physical constants that are compatible with physical theory as of that moment and mutually consistent with one another as judged by some objective criterion.

My second reason for endorsing continuation of least squares (or other) adjustments of the fundamental constants, is that a scientific field such as physics, just like an individual scientist, needs the experience, or at least the threat, of a "comprehensive examination" from time to time to keep up to snuff and in line. Without the threat of an impending new overall adjustment, experimental physicists might divide up into cliques by subfields, the members of each clique gradually acquiring a proprietary attitude toward the fundamental constant(s) in its private domain, and developing into a mutual admiration society through continual reduction of assigned uncertainties, safe from the annoyance and stimulus of discordant views and evidence from would be spoilers from the outside.

The value to a Ph.D. candidate of a comprehensive examination in his subject-matter field is, I believe, at least two-fold: He must prepare himself in depth in his specialty to be ready for the deeply probing questions of his major professor; and he must be at least conversant with the principal theoretical and experimental results in the other branches of his subject-matter field, and of the interrelations among them, in order to respond successfully to the "sharp-hooked" questions of the other members of his examining committee. So it is with a careful and conscientious adjustment of the fundamental constants: the "adjusters" individually and collectively benefit personally from the deeper and broader insight gained in making ready for the adjustment; and science as a whole benefits from the adjustment itself, and from the insights gained by its authors and reflected in their write-up.

2. Alternative Methods of Adjustment

The Method of Least-Squares is the best known and most fully developed member of a general family of "objective" adjustment procedures defined by the requirement of LEAST SUM OF Pth POWERS of ABSOLUTE RESIDUALS, $0 < P \leq \infty$.

Taking $P=2$ yields the criterion of LEAST SUM of SQUARED RESIDUALS and the associated Method of Least Squares developed by Gauss and

Legendre as an extension, to complex structured situations, of the practice of taking the arithmetic mean of a number of "equally good" measurements of a single quantity as the "best" value of the quantity afforded by the observations.

The case $P=1$ corresponds to Boscovich's criterion (1757) of LEAST SUM of ABSOLUTE RESIDUALS, whose geometrical procedure (1760) for achieving an adjustment in accordance with this criterion (and one other) was translated into algebraic form, extended, and extensively applied by Laplace in Vol. II (1799) of his *Traité de Mécanique Céleste*—see [1], e.g., pp. 434–442, 448–450, 475–478—before the advent of the Method of Least Squares. Later named by Laplace the Method of Situation, this adjustment procedure may be regarded as an extension, to complex structured situations, of "adjustment" by taking the *median* (i.e., the *middlemost* in order of magnitude) of a set of measurements of a single quantity, especially when it is desired to reduce the influence of "outliers." As in the median case, the Method of Situation, or "Least Sum" adjustment, when applied without further restriction, always leads to "adjusted values" consisting of a "chosen" subset of the actual input values, the remainder entering in to consideration "only so far as they help to *determine the choice*" ([2], p. 270). On this account, Nathaniel Bowditch expressed the view ([1], footnote to p. 434) that whereas the Method of Least Squares "is extremely well adapted" to [the reduction of] a set of observations . . . subject to the same degree of uncertainty, from the imperfections of the methods of observation, "this method, proposed by Boscovich," is to be preferred when outliers are present or there is other evidence of heterogeneity.

Letting $P\rightarrow\infty$ yields in the limit the criterion of LEAST MAXIMUM ABSOLUTE RESIDUAL, introduced by Laplace (1783), who developed procedures for effecting such adjustments, and applied them extensively in Vol. II of his *Mécanique Céleste*—see e.g., [3], pp. 417–434, 443–448, 470–474. Today, however, this approach is usually referred to as the Tchebycheff or Chebyshev Approximation—see, e.g., [3] and [4]. Such "Least Maximum . . ." adjustments may be regarded as extensions to complex structured situations, of the very early (11th century) practice of taking the mid-value between the extremes (the *midrange*) of a set of measurements of a single quantity, in order to minimize the maximum discrepancy.

Until very recently the numerical work involved in Least Sum . . . or Least Maximum . . . adjustments rendered their application intractable except in the simplest cases, but their formulation in terms of modern linear programming techniques [3–9], taken in conjunction with automatic electronic digital computers, now make these adjustments feasible in quite a variety of situations.

The Method of Least-Squares has, and I imagine always will have, definite advantages over the aforementioned alternatives: It is mathematically and numerically more tractable, though not without pitfalls [10, 11] and yields (almost automatically) measures of the statistical uncertainties (i.e., "standard errors") of the resulting adjusted values. The latter seems likely to remain an important and unique feature of Least-Squares for a long time to come. Nonetheless, my feeling is that as computer programs for carrying out Least Sum adjustments become more readily available—Least Sum adjusted values should always be obtained routinely *in addition* to Least-Squares values. If the Least-Squares and Least Sum values do not differ importantly from a practical viewpoint, then I would regard the Least Squares values as "supported," so to speak; and would stick to the Least-Squares values and their computed standard errors. If, on the other hand, the Least Sum and Least Squares adjusted values differ to a practically important extent, then I would take this to be an indication that the input data require further scrutiny and maybe a bit of cleansing before proceeding further.

For instance, as was Laplace's practice, it may be a good plan to carry out the Least Maximum Residual adjustment, take a look at the resulting maximum residual, and decide whether or not it is unbelievably too large in relation to the present state of the art.

3. Distinct Values and Components of Imprecision

Lest my endorsement of continuation of the practice of Least Squares (or other) adjustments of the fundamental constants leave with you the impression that I always favor taking an average among two or more available values, I shall now consider an instance where I feel that "averaging" of any kind is completely unwarranted; and three additional cases in at least one of which averaging seems called for, resulting in a mean with a somewhat (and sometimes considerably) larger standard error than is commonly assigned.

3.1. Nitrogen Content of SRM 1093

I refer to the situation here at the NBS with respect to the nitrogen content of Standard Reference Material 1093 (SRM 1093), a "valve steel." SRM 1093 has been analyzed repeatedly for nitrogen content by the NBS Office of Standard Reference Materials, using two "standard" methods of analysis, Vacuum Fusion and a Pressure Bomb-Distillation-Titrimetric method that is a recent modification of the "classical" Kjeldahl method of analysis for nitrogen. The nitrogen content obtained from about 50 determinations by the Vacuum Fusion method is reported [12] *as* 3640 ± 76 ppm; from about 8 determinations by the Pressure Bomb-Distillation-Titrimetric method, *as* 4807 ± 10 ppm. The "76" and "10" are in each case the standard deviations of the individual determinations involved (*not* of the respective means), and reflect both analytical error

and heterogeneity of the material. The two distributions of results do not even come close to overlapping—they are miles apart.

Although both methods aim to determine the same "target value," i.e., *the* nitrogen content of SRM 1093, it is obvious that there is a well determined systematic difference between results obtained by the two methods; either one, or the other, or both of the methods is (are) affected by a substantial systematic error. Work is now underway at the NBS to discover and resolve the systematic error(s) existing in either or both of these methods. In the meanwhile, it is clear that 3640 ppm should be taken as *the* nitrogen content of SRM 1093 when a sample of SRM 1093 is used to provide a "bench mark" in nitrogen analyses of other steels by the Vacuum Fusion method; and that 4807 ppm shall be taken as *the* nitrogen content, when samples of SRM 1093 are used in nitrogen analyses of other steels by the Pressure Bomb-Distillation-Titrimetric method. To do otherwise, for example, to "split the difference" and use the mid-value 4223.5 ppm in applications of both methods would almost certainly introduce a substantial hidden systematic error into the results of one, the other, or both—and reduce the incentive to resolve the difference!

3.2. Acceleration of Gravity Determinations at Ottawa

A case of an entirely different sort is presented by the August 1958 and December 1959 data obtained in the absolute measurement of the acceleration due to gravity at Ottawa by Preston-Thomas *et al.* [13] and portrayed in figure 1, which is a facsimile reproduction of their figure 9. Although the upper two histograms in figure 1 overlap, it is quite evident without conducting a formal statistical test that the difference between their means is "statistically significant." This is the same conclusion to which we came in the case of the two sets of SRM 1093 nitrogen-content data, i.e., "the two sets of data are different," but our interpretation of the difference and our resulting action should in the present instance, I feel, be quite different.

The observed difference between the "locations" of these two histograms must be ascribed either to some (unsuspected) difference between Rule 1 and Rule 2, or to some other difference between the 1958 and the 1959 "Drops," or both. A glance at the bottom histogram showing the results of the 1957 preliminary drops with Rule 1, reveals that, in spite of the much greater dispersion of these preliminary drops, their mean is remarkably close to that of the 1958 drops with Rule 1, suggesting that the observed difference between the 1958 and 1959 drops is due principally to some difference between Rules 1 and 2. Furthermore, if we regard these two rules as "samples" from a (conceptual) "population" of rules made "exactly like Rules 1 and 2," and if the desired difference is really an expression *of* rule-to-rule variation, then we can say that the chances are about even (i.e., about one-third) that the mean of 32 drops with a new Rule 3 would fall below the (1959) mean of the 32 drops with Rule 2, or above the (1958) mean of the 32 drops with Rule 1, or between these two means. (Cf, e.g., [14], p. 144.)

If we thus regard the observed difference between the means of the top two histograms as being due to the *sampling variation* of some uncontrolled property of Rules made "exactly like" Rules 1 and 2, then the simplest unbiased (and least assumption-dependent) course of action is to take the simple *unweighted* arithmetic mean of the *means* of the 32 drops with Rules 1 and 2 respectively, and assign to the resulting value, 980.61315 cm s^{-2} a standard deviation equal to the square root of one-quarter of the squared difference, i.e., one-half of the

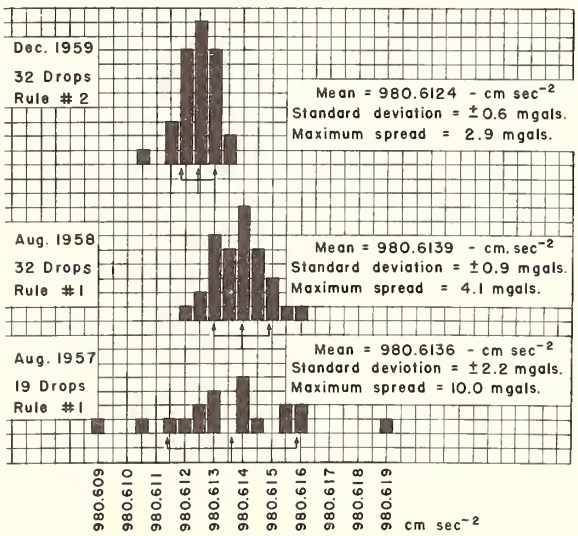


FIGURE 1. *Histograms of values of acceleration of gravity, g, from drops with two rules (Reproduced with permission from Preston-Thomas et al. from [13].*

observed absolute difference; namely, 0.00075 cm s^{-2} or 0.75 mgal. In so doing, however, one must not lose sight of the fact that this computed standard deviation is very loosely determined indeed, being based (in statistical language) on only "a single degree of freedom," so that 95 percent-confidence limits for the "true value" of this standard deviation range from about 0.27 mgal to about 13.5 mgal (NBS Handbook 91, Table A-20); and 95 percent-confidence limits for the "limiting mean" ([15], p. 103; [16], pp. 168–169) of which the observed mean is but a single "determination" are 980.61315 ± 0.00953 cm s^{-2}. This enormously wide band of uncertainty can be reduced only by either averaging out rule-to-rule variation by employing more rules, or by discovering the cause of such rule-to-rule variation and eliminating it.

3.3. Heyl's Determinations of the Constant of Gravitation, G

All of you, I am sure, are familiar, or at least, acquainted with the "redetermination of the constant of gravitation," G, carried out at the National Bureau of Standards in the mid-1920's by the late Paul R. Heyl by means of a torsion balance, with small moving masses consisting of a pair of gold balls, a pair of lacquer-coated platinum balls, and a pair of balls made of optical glass. The mass of each of the balls was about 50 g. The gold and platinum balls were of about the same size, inasmuch as the densities of gold and platinum differ by only about 10 percent, but, the density of the optical glass being between 1/5 and 1/6 of the densities of gold or platinum, the glass balls were correspondingly larger. In the ensuing formal publication [18], Heyl stated (p. 1251): "This use of different materials was not prompted by any suspicion that a specific difference in attraction might be found. Rather was it the result of circumstances." The original plan called for the use of gold balls, by analogy with the earlier (1895) work of C. V. Boys, and Carl Braun's use (1897) of gilded brass balls; "but in the present work it was found on opening the container, after a period of several months required for the observations, that the gold balls had absorbed quite appreciable quantities of mercury, probably derived in vapor form from the manometer connected to the container." "A second set of instruments was made with platinum balls, coated thinly with lacquer . . .. No appreciable change of weight was observed in this case." For good measure, "a third set was made with balls of optical glass, the idea being that holes in the interior could be detected visually."

Results of these three sets of experiments are shown in figure 2, the plotted points for gold being the observed values after correction for the increase of mass of the gold balls from absorption of mercury.

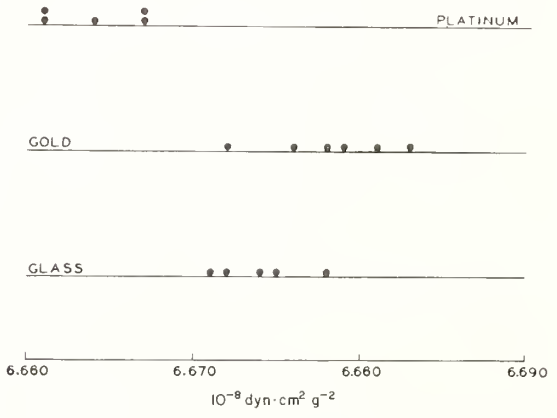


FIGURE 2. *Plots of Heyl's torsion-balance measurements of the gravitational constant, G, using small masses of three different materials.*
(Reproduced with permission from Shewhart [28], p. 69. Copyright 1939, The Graduate School, U. S. Department of Agriculture.)

It is evident from this figure, without formal statistical analysis, that systematic differences exist between the results of these three materials. From the three means and their respective standard errors:

Platinum:	6.6640 ± 0.0013
Glass:	6.6740 ± 0.0012
Gold:	6.6782 ± 0.0016

it is clear that the differences between each mean and the other two are "statistically significant." As Heyl remarked some years later ([19], p. 1), "the results of each material were reasonably concordant among themselves but there appeared to be considerable difference with the nature of the material."

Here we have a situation not unlike that of the Canadian drops with Rule 1 and Rule 2, but with one essential difference: not only were three (rather than just two) pairs of balls used, but in addition, the three pairs were made with three quite different materials. (I wonder what would have been the outcome if the Canadians had used three Rules, or better three *pairs* of Rules, of different materials.) What should be done? In the absence of further research satisfactorily accounting for, and making it possible to explain away the observed differences of these sets of measurements, I feel that we have no alternative to inferring that the observed differences reflect some unexplained sensitivity of such torsion-balance experiments to the materials actually used for the "small masses"; that we must, therefore, regard these three particular materials as merely a "sample of three" from the "universe" or "population" of materials that might have been used, or might be used in repetitions of these experiments; and that, in consequence, the simple unweighted arithmetic mean of the three means should be taken, and its statistical standard deviation estimated statistically from the dispersion of these three means, as indicated in section 3.5. (What would I do, you may ask, if instead of results yielded by a single experimental procedure using three different materials, the results of figure 2 were the outcomes of procedures based on three distinctly different experimental methods? My answer would depend on the availability or nonavailability of certain other relevant information; and if available, on what its message was.)

Heyl remarked ([18], p. 1286) that "it is evident that the gold ball results are much less reliable than either of the other sets." For this reason, and perhaps others, he took a weighted mean, retaining the gold results at one-third the weight of either platinum or glass. I regard such weighting as indefensible in view of the large set-to-set differences, even if the greater scatter of the gold ball results be interpreted to mean that they are "less reliable." And I do not consider the evidence convincing that they are "less reliable": the squared standard deviations of the individual measurements of the platinum, gold, and glass sets are 9, 15, and 7.5, respectively, in

units of the third decimal; the ratio of the largest to the smallest is only 2, a ratio that has about a 50:50 chance of being exceeded when comparing random pairs of sets containing only five or six measurements, and more likely to be exceeded than not when comparing the largest and smallest two out of three. It would be instructive to experimenters to spend a little time playing around with sets of synthetic measurements (i.e., from a table of "random normal deviates") they would soon learn to place much less reliance than they commonly do on standard deviations calculated from small sets of data.

As a postscript to all of this, when Heyl and Peter Chrzanowski carried out a "new determination of the constant of gravitation" [19], they avoided gold and glass balls, and were content to use only a pair of platinum balls weighing about 87 g each; but, alas, this time they used two different types of filament, "hard-drawn" and "annealed," in the torsion balance itself. Among their concluding remarks we find: "It will also be noted that the present results show the same peculiarity of pattern as did those in the 1930 paper. The results with the hard-drawn filament differ from those with the annealed filament by an amount greater than can be accounted for by the departures from the mean" ([19]), p. 31).

3.4. The NPL Determinations of the Ampere

A somewhat but, perhaps, not exactly similar situation is evidenced by Vigoureux's 1962 and 1963 measurements of current by the NPL Current Balance portrayed in figure 3, which is a copy of his figure 2 [20]. At first glance, these data seem to present essentially the same problem of analysis and interpretation as the Canadian acceleration of gravity data, except in that instance the two sets of data correspond, respectively, to measurements taken before and after Rule 1 was removed and replaced by another nominally identical and presumably equivalent Rule 2, whereas in the present instance the two sets of data correspond to measurements taken before and after the very same coils were removed and then restored.

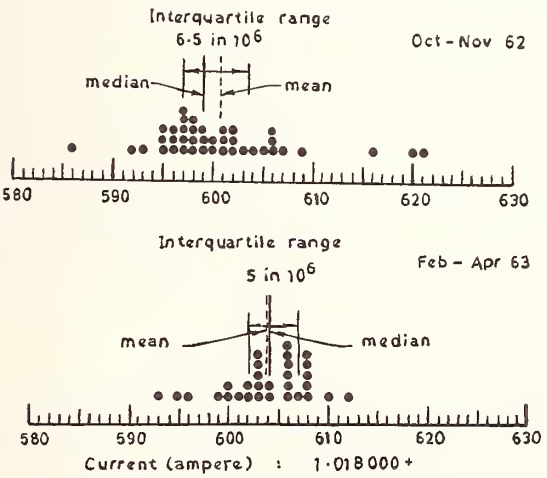


FIGURE 3. *Histograms of 40 measurements of current by NPL current balance in 1962 before removal of coils, and of 30 measurements in 1963 after coils were restored.*
(Reproduced with permission from Vigoureux [20]. Crown copyright reserved.

By analogy with the Canadian experiment, we may be inclined in the present instance to attribute the shift in the mean to some uncontrolled "cause" associated with removal and restoration of the coils; presume that, if the coils were removed and restored again, a fresh shift of the mean (in either direction) would ensue; and carry out an analysis along the lines of that above for the two sets of acceleration-of-gravity data, obtaining 1.018 6023 A as the simple (unweighted) arithmetic mean of the two means, with an assigned standard deviation of 0.0000016 A (or 1.6 ppm), the latter figure again being very ill-determined from only one degree of freedom.

On the other hand, there is an important difference between the acceleration-of-gravity experiment and this ampere-measurement experiment: The Canadians did *not* try a third rule to see what effect that would have, but Vigoureux *did* "disturb" his coils to see whether he could eliminate, or at least vary the effect. He wrote:

> "As soon as the first few measurements of the second series indicated the trend, efforts were made to discover the reason for the shift of the mean. Several times the suspended coils were disturbed and relevelled, and the positions of the large coils were adjusted anew, but the positions repeated remarkably well. This precaution had not been taken for the first series . . .". ([20] p. 5).

Hence, one is led to conjecture that at least some of the measurements of the first set were affected by a systematic error, or, perhaps, by systematic errors, stemming from a common cause that was permanently eliminated when the coils were removed.

Some light might be thrown on this issue if the time-order of measurement of the points plotted in the two plots of figure 3 were available. If such time ordering of the measurements revealed a pattern of variation that strongly suggested (a) the presence of an unidentified disturbance apparently affecting the measurements either intermittently, or throughout some major segment(s) of the first set, and (b), *no* similar effect in the second set, then it would seem advisable to discard the first set entirely, and base the further analysis on the mean 1.018 603 87 A of the second set only; and on the computed standard deviation of this mean, i.e., 0.000 000 80 A which, being based on 29 degrees of freedom, is actually quite well determined (95%-confidence limits: 0.63 to 1.06 ppm). If, on the other hand, there were evidence of heterogeneity in the second set also, and the 70 measurements *considered in time order* seemed to break up into at least two more or less distinct sets, then an analysis along the lines outlined in section 3.5 would be appropriate. Otherwise, in the absence of such enlightening information, it is safer, and certainly more conservative, to proceed

in the same manner as in the case of the acceleration-of-gravity data—which is about what Vigoureux did.

3.5. Unweighted Mean Recommended When Components of Variance Are Present

When "systematic differences" among subsets of measurements of a single quantity are evident or revealed by statistical analysis, use of the unweighted mean of the subset means is recommended, rather than a weighted mean of these means. Specifically, when a set of N measurements of a single quantity consists of k (≥ 2) subsets of $n_1, n_2, \cdots, n_k (\sum_1^k n_i = N)$ measurements, respectively, made on each of k different "occasions," or corresponding to each of k different "values" of some qualitative factor (e.g., different Rules, balls of different materials), and these "occasions" or "values" may be regarded as a sampling of the variations of procedure, apparatus, environmental conditions, etc., allowable in "repeated applications" of what will be considered to be the "same" method of measurement applied to the measurement of the same quantity,[1] and if the data indicate, or at least suggest, the presence of systematic differences between the subsets, then the simplest unbiased (and least assumption dependent) course of action is to take the unweighted arithmetic mean,

$$\bar{X}_{i.} = \sum_1^k X_{i.}/k, \tag{1}$$

of the individual subset means,

$$X_{i.} = \sum_1^{n_i} X_{ij}/n_i \qquad (i = 1, 2, \cdots, k), \tag{2}$$

as the value of the quantity measured indicated by the data in hand; and ascribe to $\langle X_{i.} \rangle$ a standard error given by the square root of

$$\sum_1^k (X_{i.} - \bar{X}_{i.})^2/k(k-1), \tag{3}$$

with $(k-1)$ degrees of freedom.

Explanation. The basis for this recommendation is as follows: Under the above circumstances the jth measurement of the ith subset, X_{ij}, may be written in the form

$$X_{ij} = \xi + D_i + E_{ij},$$
$$(i = 1, 2, \cdots, k;\ j = 1, 2, \cdots, n_i), \tag{4}$$

where ξ denotes the quantity measured, an unknown constant; D_i, a fixed displacement or systematic error peculiar to, and affecting all measurements of, the ith subset, the D_i collectively being regarded as the experienced values of random variables D_i, $(i = 1, 2, \cdots, k)$, independently and identically distributed about a "population" mean of zero, with a "population" standard deviation σ_D; and the E_{ij}, measurement "errors" that for *fixed* i may be regarded as independent identically distributed random variables with a "population" mean of zero, and a "population" standard deviation of σ_{E_i}, $(i = 1, 2, \cdots, k)$, possibly different from set to set, and independent of the D_i and of the E_{hj}, $(i \neq h = 1, 2, \cdots, k;\ j = 1, 2, \cdots, n_h)$. When the entire set of N measurements may be thus characterized, the *variance* (i.e., squared standard deviation) of the mean, $X_{i.}$, of the ith subset *as a determination of* ξ is

$$V\{X_{i.}\} = \sigma_D^2 + (\sigma_{E_i}^2/n_i) \qquad (i = 1, 2, \cdots, k), \tag{5}$$

so that the variance of $\bar{X}_{i.}$, the mean of the subset means, *as a determination of* ξ is

$$V\{X_{i.}\} = (1/k)\{\sigma_D^2 + (1/k) \sum_1^k (\sigma_{E_i}^2/n_i)\}, \tag{6}$$

and the *expected*, or average, value of (3) in a series of repeated experiments of the same structure, is exactly equal to the right-hand side of (6). Hence the square root of "observed" value of (3) can be used directly as an estimate of the standard error of $\bar{X}_{i.}$ without first computing estimates, $\hat{\sigma}_d^2$ and $\hat{\sigma}_{E_i}^2$, $(i = 1, 2, \cdots, k)$, of the quantities on the right-hand side of (5). This practical convenience is not shared by any other mean of the $X_{i.}$; and, in particular, not by the grand mean,

$$X_{..} = \sum_1^k n_i X_{i.} / \sum_1^k n_i,$$

except when the subset sizes are all equal, $n_i = n$, *and* the subset measurement errors all have the same standard deviation, $\sigma_{E_i} = \sigma_E$, say.

At this juncture persons familiar with the method of least squares may be wondering why I have not thus far endorsed use of the "optimally" weighted mean,

$$\hat{\xi} = \sum_1^k w_i X_i / \sum_1^k w_i,$$

say, with weights, w_i, proportional to the reciprocals of the right-hand side of (5) for $i = 1, 2, \cdots, k$, respectively, which certainly must be more precise, i.e., have a smaller standard error. There are three reasons: First, when the actual values of the variances σ_D^2 and $\sigma_{E_i}^2$, $(i = 1, 2, \cdots, k)$, are known, which is rarely, if ever, the case outside of "textbook" examples. Cochran's collection of numerical examples ([21], Table III, p. 114) shows that when the ratio

$$\sigma_D^2 \Big/ \frac{1}{k} \sum_1^k \frac{\sigma_{E_i}^2}{n_i} \tag{7}$$

exceeds 2, which is very common in practice, especially when $n_i > 2$ for all i, the standard error of the unweighted mean, $\bar{X}_{i.}$, is unlikely to exceed the standard error of the "optimal" mean, $\hat{\xi}$, by as much as 4 percent; and that when (7) lies between 1 and 2, the excess will rarely be as much as 15 percent. The "traditional" weights,

$$n_i/\sigma_{E_i}^2 = w_i', \text{ say,} \tag{8}$$

[1] For a detailed discussion of the concept of a "repetition" of a measurement by the "same" measurement process, or "same" method of measurement, see [16], pp. 181–183.

are optimal only when $\sigma_D^2 = 0$, and the corresponding weighted mean, $\bar{X}_{i.}'$, say, has a smaller standard error than the unweighted mean $\bar{X}_{i.}$, only when (7) is in the neighborhood of $\frac{1}{2}$ or less.

Second, when the value of σ_D^2 is unknown, the usual case, a value for it must be deduced from the data in hand. Unfortunately, the customary "analysis-of-variance" procedures for estimating "components of variance" do not yield "optimal" estimates of σ_D^2, *except* when the n_i are all equal ([22], pp. 347–348, 351–354)—an uncommon situation in experimental physics—so that one must resort to iterative solutions such as have been indicated by Cochran ([21], p. 113) and treated in detail by Mandel and Paule [23]. Tukey, in a related study [24], observed (p. 43): "Weighting [subset] means equally seems to give a better estimate than classical proportional weighting (i.e., weighting by n_i) for the between variance component as soon as (i) the between component exceeds $\frac{1}{2}$ of the within component in a moderately unbalanced design (i.e., moderate inequalities among the n_i), or (ii) the between component exceeds the within component in a substantially unbalanced design. Slight further gains come from intermediate weighting." Which leads us back to our (3) as a good "start".

Third, Cochran points out ([21], p. 115) that "in practice the situation is much more favorable to the unweighted mean," $\bar{X}_{i.}$ than is implied by his numerical results cited above for the case where the values of σ_D^2 and $\sigma_{E_i}^2$ are known. Their estimation from the data in hand results in an increase in the standard error of the otherwise "optimal" mean, ξ, to which there is no corresponding increase in the standard error of the unweighted mean, $\bar{X}_{i.}$. Consequently, when the n_i are all small, say <16, and unequal, the actual standard error of the unweighted mean $\bar{X}_{i.}$ will generally be less than that of the otherwise "optimal" mean, $\hat{\xi}$, except when $\sigma_D^2 = 0$, at least to a good approximation—even though this may not be revealed by their respective standard errors computed from the data.

Furthermore, it needs to be pointed out that considerable under-estimation of the relevant standard error can result when $N = \sum_1^k n_i$ measurements such as we have been discussing are lumped together and analyzed without taking into account the fact that they really consist of k distinct subsets of *unequal* sizes. In such an event, one would almost certainly use the grand mean, $X_{..}$, and take the square root of

$$\sum_1^k \sum_1^{n_i} (X_{ij} - X_{..})^2 / N(N-1) \qquad (9)$$

as an estimate of the standard error. Brownlee points out ([25], pp. 323–324) that, when $\sigma_{E_i} = \sigma_E$ and $n_i = n$, $(i = 1, 2, \cdots, k)$, the expected value of (9) is

$$(kn)^{-1}\{[n(k-1)/(kn-1)]\sigma_D^2 + \sigma_E^2\} \qquad (10)$$

whereas under these circumstances the squared standard error of $X_{..}$ is actually

$$(kn)^{-1}\{n\sigma_D^2 + \sigma_E^2\}. \qquad (11)$$

The former is clearly always less than the latter, *except* when $\sigma_D^2 = 0$, when the kn measurements do constitute a single homogeneous set, *or* when $n = 1$, complete heterogeneity. As $\sigma_D^2/\sigma_E^2 \to \infty$, their ratio tends to $(k-1)/(kn-1)$, which for n fixed tends to $1/n$ as $k \to \infty$; and for k fixed, tends to 0 as $n \to \infty$. In particular, when $\sigma_D = 2\sigma_E$ and $n = 2$, the ratio is always less than 5/9 $(k = \infty)$ and decreases to 0.407 as $k \to 2$. When $\sigma_D = 2\sigma_E$ and $n = 10$, the ratio is always less than $5/41 \simeq 1/8$ and decreases to 0.076 as $k \to 2$. The "story" is the same, although much more complex, when the n_i or the σ_{E_i} are unequal. "Thus a very serious error may be made by assuming that we have *kn independent* observations when in point of fact we do not" ([25], p. 324).

4. Concluding Remarks

A few remarks are now in order on the traditional procedure for arriving at a measure of the "total uncertainty" of an experimental result as a determination of a particular target quantity of interest, and on the combination of the information provided by two or more experimental results presumed to be determinations of the same quantity, in the light of the preceding examples and discussion.

The traditional procedure consists, as you know, of taking the "total uncertainty" of an experimental result to be equal to the square root of the sum of the squares of (i) a measure of its "statistical uncertainty," usually the value of its *standard deviation*, or *probable error*, computed from the deviations of the individual measurements of which it is a "mean," and (ii), subjective estimates of additional "uncertainties" corresponding to various explicitly recognized sources of potential systematic error, expressed in standard deviation or probable error form (usually without any indication of the rule-of-thumb or other means by which this was achieved). As you know also, the current and historical literature on the fundamental constants are rife with examples of two or more alleged determinations of some particular "constant" the assigned total uncertainties of which do not even overlap—eloquent testimonials to the failure of this traditional procedure in actual practice. And I have heard at least one speaker at this Conference state that he would not be especially upset if his contribution to the melting pot were found to be off by 4 or 5 times its assigned total uncertainty. To a statistician or a mathematician—if I may so style myself even though my undergraduate major was mathematical physics—this situation is scandalous! What is wrong? What needs to be done to bring the situation into line?

The primary source of the trouble consists, I believe, in gross underestimation of the statistical uncertainty of the reported mean. Part of this underestimation stems from the use of computational

(and weighting) procedures to derive a computed value for the standard deviation of the mean that fail to make proper allowance, if any, for a "between subsets" component of variance—and in the process imply that the mean involved is based on far more *independent* measurements than it really is. (The number of *independent* measurements embodied in the mean that constitutes "the result" of the Canadian g experiments is only 2, not 64; Heyl's 1930 value for G is a "mean" of only 3 "independent" values; and his 1942 value, of only 2.) In consequence, the resulting "standard errors of the mean" are not only gross underestimates, but also are regarded as being far more precisely determined than they actually are. Even when the standard error of the mean is properly determined the resulting value is liable to be treated as exact in arriving at a value for the "total uncertainty" via the traditional formula, whereas, when derived from only 3 degrees of freedom, say, the chances are about 1 in 10 that it will underestimate the "true" standard error of the mean by a factor of 2.25 or more; and from only 1 degree of freedom, by a factor of 8 or so. The additional uncertainty from this source would be revealed if, instead of reporting the "total uncertainty" in standard deviation or probable error terms, the total uncertainty were expressed in "units" corresponding to a normal deviate of "2σ," say. To this end, the subjectively estimated uncertainties would merely need to be multiplied by 2 or 3, according as they were initially expressed as "standard deviations" or "probable errors"; and the computed statistical standard error of the mean inflated by the "2σ" factors given in Joiner's table of "Student-t deviate corresponding to a given normal deviate" [26].

A more serious, and less easily remedied, source of underestimation of the statistical uncertainty of the mean is the failure of most experimenters to adequately sample the variations of procedure, apparatus, environmental conditions, etc., that would in practice inevitably occur in any realistic repetition of the "same" experiment, if the apparatus and data sheets were totally destroyed and a fresh start were made from the beginning. After all, an important, and, perhaps, the main purpose of a measure of statistical uncertainty of the "mean" is to provide an indication of how closely the result of this particular experiment is likely to agree with other results that might have been obtained in this instance, or might be obtained if the entire experiment were carried out afresh "from the beginning."

The message of the experiences of Drs. Heyl and Preston-Thomas and their associates is loud and clear: When experimenters change parts of their apparatus, results obtained change also. And Dr. Vigoureux's experience reveals that similar consequences can ensue when a portion of the apparatus is disassembled and then reassembled. A corollary is that if an experimenter fails to include such changes of essential parts of his apparatus in his experimental program then he will have no opportunity, and be saved the embarrassment of discovering the effects of such changes; but he will also almost certainly underestimate the standard error of his mean, often considerably, and thus add to the woes of those who wish to compare his final results with those of others, or combine them to form "an adjusted" value. To this the experimenter will no doubt reply that inasmuch as there are so many things that could be changed, even if only a few are changed the number of possible combinations mounts astronomically. Dr. W. J. Youden has responded to this latter point, by showing how to select small strategic subsets of all possible combinations, in a paper published nearly a decade ago in *Physics Today* [27], and again in his retiring address as President (1968) of the Philosophical Society of Washington, entitled "Enduring Values," as yet regrettably unpublished.

I realize that some of the subjectively estimated uncertainties tabulated by experimenters are intended, when combined "in quadrature" with the statistical standard error of the mean, to provide an inflated "standard error" of the mean that can be used to "provide an indication of how closely the result of this experiment is likely to agree with other results that—might be obtained if the entire experiment were carried out afresh 'from the beginning,' " but experience shows that this is a delusion: there is no substitute for comprehensive realistic repetitions of the entire experiment!

Which brings me to Barry Taylor's remark in this Panel Discussion that no cry of anguish is heard from the scientific community at large when previously accepted values of the fundamental constants are changed by 4 to 5 times their previously assigned total uncertainties. This disturbs me because it supports my long-time belief that least-squares adjustments of the fundamental constants belong to a dream world, the uncertainties assigned to the respective input data being almost certainly too small as a result of inadequate replication of the basic experiments involved, and this shrinkage of uncertainties compounded in those ascribed to the final adjusted values as a result of employing weighting schemes that the very data themselves often indicate (to an experienced professional statistician's eye) to be inappropriate. Moreover, my own experience with very similar situations suggests that, corrective action is not likely to be taken until the inherent discrepancies involved in the present approach really "make a difference" in some practically important situation, i.e., not until some applied scientist or engineer *really* required not only the "last few decimal places," but also realistic and dependable assessments of their uncertainties, to safely avoid potentially disastrous consequences. The late Walter A. Shewhart put the matter eloquently thus [28]:

> "... Thus we see why it is that the applied scientist cannot stop with *making* estimates of precision and accuracy—he must also *act* on the basis of such estimates. He knows that *this action will reveal his mistakes*, and what is more important, he knows that such mistakes may carry with them serious consequences.
>
> "The applied scientist in order to be 'successful' cannot afford to make too many mistakes even though they be

small, and in no case can he afford to make a mistake that is large enough to cause serious trouble. His tendency is to be cautious in accepting any estimate of precision or accuracy as a basis for action. In his language, he wants to be 'sure' of his estimates before making them the basis of mass production practices. He does not consider his job simply doing the best he can with the available data; it is his job to *get enough data before making his estimates.*"

Until the day comes when experimenters provide the "adjusters" with the grist that they really need for their adjustment mill, I recommend that the adjusters discard individual determinations of a particular quantity very sparingly, preferably on non-statistical grounds, and certainly not on the basis of the "chi-square" test alone when applied to results obtained by much the same method—a "significant" chi-square in such cases indicates to me simply the presence of a "between" component of variance that must be taken into account. When the results compared correspond to radically different methods, the situation is a bit sticky, unless the values fall into two or more recognizable subgroups, one of which can be chosen on "other" grounds. When there is no basis for choice, then, in the absence of a basis for obtaining a realistic measure of the statistical uncertainties involved, it seems to me appropriate to proceed as recommended in section 3.5. The resulting mean will no doubt tend to be assigned a standard error that is too large—but this unsatisfactory state of affairs may persuade some experimenters to run their experiments again, to help provide more realistic measures of the uncertainties actually involved. Present practice penalizes the realistic experimenter who explores the effects of various changes in apparatus, procedure, and passage of time, and winds up with an honest but comparatively large standard error of his mean, in favor of the naive experimenter who holds everything as constant as possible, and comes up with a small but meaningless standard error of his mean. A realistic approach to the problem of assigning "best" values to the fundamental physical constants with valid indications of their standard errors requires that we stop assuming values for components of error that we have never measured and, through experiments of adequate scope, find out what the facts actually are. A corollary is that, when some experimenter has uncovered an unsuspected source of variation—by changing rules, by using balls or filaments of different materials, by disassembly and reassembly of a portion of his apparatus, and so forth—and has obtained a measure of the effect on his final result of variations in this factor, his findings should not be ignored, but rather be fully and explicitly taken into account not only in combining his and previous results by essentially the same method, to obtain a new "best" value, but also in the design, execution, and interpretation of subsequent experiments of the same kind.

5. Acknowledgments

Finally, I want to thank J. M. Cameron, H. H. Ku, and J. Mandel of the NBS for a number of helpful suggestions that have clarified and strengthened a number of points stated somewhat vaguely in an earlier draft and in my oral presentation.

6. References

[1] de Laplace, P. S., Celestial Mechanics, Vol. II, translated with notes and commentary, by Nathaniel Bowditch, pp. 417–486 (Hilliard, Gray, Little, and Wilkins, Publishers, Boston, Mass., 1832; facsimile reproduction, Chelsea Publishing Company, Bronx, N.Y., 1966).

[2] Gauss, K. F., Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections, translated, with an Appendix, by Charles Henry Davis, p. 270 (Little, Brown and Company, Boston, Mass., 1857; facsimile reproduction, Dover Publications, Inc., New York, N. Y., 1963).

[3] Goldstein, A. A., and Cheney, W., A finite algorithm for the solution of consistent linear equations and inequalities, and for the Tchebycheff approximation of inconsistent linear equations, Pacific Journal of Mathematics, **8**, No. 3, 417–427 (Fall 1958).

[4] Stiefel, E. L., Note on Jordan elimination, linear programming, and Tchebycheff approximation, Numerische Mathematik **2**, No. 1, 1–17 (1960).

[5] Stiefel, E. L., Numerical methods of Tchebycheff approximation, in On Numerical Approximation, Ed. R. E. Langer (University of Wisconsin Press, Madison, Wis., 1959), pp. 217–232.

[6] Goldstein, A. A., Levine, N., and Hereshoff, J. B., On the "best" and "least *Q*th" approximation of an overdetermined system of linear equations," Journal of the Association for Computing Machinery, **4**, No. 3, 341–347 (July 1957).

[7] Wagner, H. M., Linear programming techniques for regression analysis, Journal of the American Statistical Association, **54**, 206–212 (Mar. 1959).

[8] Rice, J. R., The Approximation of Functions, Vol. **1**, Linear Theory, pp. 112–116 (Addison-Wesley Publishing Company, Inc., Reading, Mass., 1964).

[9] Davies, M., Linear approximation using the criterion of least total sum of deviations, Journal of the Royal Statistical Society, Series B, **29**, No. 1, 101–109 (1967); and No. 3, 587 (1967).

[10] Wampler, R. H., An evaluation of linear least squares computer programs, J. Res. Nat. Bur. Stand. (U.S.), **73B**, (Math. Sci.), No. 2, 59–90 (April–June 1969).

[11] Golub, H., Matrix decomposition and statistical calculations, in Statistical Computations, Eds. R. C. Milton and J. A. Nelder, (Academic Press, New York, N. Y., 1969), pp. 365–397.

[12] Certificate of Analysis: Standard Reference Material 1093 (National Bureau of Standards, Washington, D.C., March 4, 1969).

[13] Preston-Thomas, H., Turnbull, L. G., Green, E., Daufhinee, T. M., and Kalra, S. N., An absolute determination of the acceleration due to gravity, Canadian Journal of Physics, **38**, No. 6, 824–852 (June 1960).

[14] Youden, W. J., Sets of three measurements, Scientific Monthly, **77**, No. 3, 143–147 (Sept. 1953).

[15] Dorsey, N. E., and Eisenhart, C., On absolute measurement, Scientific Monthly **77**, No. 2, 103–109 (Aug. 1953); reprinted in [17].

[16] Eisenhart, C., Realistic evaluation of the precision and accuracy of instrument calibration systems, J. Res. Nat. Bur. Stand. (U.S.), **67C**, No. 2, 161–187 (April–June 1963); reprinted, with corrections, in [17].

[17] Ku, H. H., Editor, Precision Measurement and Calibration. Selected NBS Papers on Statistical Concepts and Procedures, Nat. Bur. Stand. (U.S.), Spec. Publ. 300, Vol. **1**, 436 pages (U.S. Government Printing Office, Washington, D.C., Feb. 1969).

[18] Heyl, P. R., A redetermination of the constant of gravitation, Bur. Stand. (U.S.) J. Res. **5**, 1244–1290 (RP 256, Dec. 1930).

[19] Heyl, P. R., and Chrzanowski, P., A new determination of the constant of gravitation, J. Res. Nat. Bur. Stand. (U.S.), **29**, 1–31 (RP 1480, July 1942).
[20] Vigoureux, P., A determination of the ampere, Metrologia **1**, No. 1, 3–7 (Jan. 1965).
[21] Cochran, W. G., Problems arising in the analysis of a series of similar experiments, Supplement to the Journal of the Royal Statistical Society **4**, No. 1, 102–118 (1937).
[22] Graybill, F. A., An Introduction to Linear Statistical Models, Volume **1**, pp. 337–347, 351–354 (McGraw-Hill Book Company, Inc., New York, N.Y., 1961).
[23] Mandel, J., and Paule, R. C., Interlaboratory evaluation of a material with unequal numbers of replicates, Anal. Chem **42**, 1194–1197 (1970).
[24] Tukey, J. W., Variances of variance components: II. The unbalanced single classification, Annals of Mathematical Statistics **28**, No. 1, 43–56 (March 1957).
[25] Brownlee, K. A., Statistical Theory and Methodology in Science and Engineering, 2nd edition, pp. 318–330 (John Wiley & Sons, Inc., New York, N.Y., 1965).
[26] Joiner, B. L., Student-*t* deviate corresponding to a given normal deviate, J. Res. Nat. Bur. Stand. (U.S.), **73C**, (Eng. and Instr.), Nos. 1 and 2, 15–16 (January–June 1969).
[27] Youden, W. J., Systematic errors in physical constants, Physics Today **14**, No. 9, 32–34, 36, 38, 40, 43 (Sept. 1961). Reprinted in [17].
[28] Shewhart, W. A., Statistical Method from the Viewpoint of Quality Control, Ed. W. E. Deming, p. 121 (The Graduate School, The Department of Agriculture, Washington, D.C., 1939).

PANEL DISCUSSION

D. N. LANGENBERG: If I may join in the assault on our chairman, I feel compelled to comment that his anonymous four-high and two-low distribution looks suspiciously like the well-known example of μ_p/μ_n.

P. L. BENDER: I didn't do that. (*Laughter*).

D. N. LANGENBERG: In any case, in defense of Barry Taylor's point of view about the desirability of a certain dramatization of discrepancies, I think I ought to reiterate what Barry said a couple of days ago. That is, since it's now abundantly clear (I would say almost certain) that we picked the wrong hump of the camel in that μ_p/μ_n distribution, this little vestpocket drama now points the finger of suspicion rather clearly at the single measurement of the Faraday which exists, and I think points up the importance of further measurement of the Faraday and also some of the x-ray measurements of lattice constants which also give information about the Avogadro number.

Finally, a question to the panel as a whole: Would any of you like to comment on the desirability of having not one but two sets of constants, a safe, low-resolution set which could be published in all the requisite places with the official seal of all the requisite committees, and a dangerous, high-resolution set which we could perhaps pass around surreptitiously in some kind of underground publication? (*Laughter*).

P. L. BENDER: Would anyone care to comment?

E. R. COHEN: I have a comment to make, taking the opportunity—not in answer to Don's specific question—(*Laughter*)—but after all, we have now both been caught. DuMond and Cohen made the unfortunate decision that if they had two pieces of observation that were discrepant, one of which was infinitely accurate but with an uncertain theory, and another one which seemed to be reasonably good with a reasonable accuracy, you ought to pick the experiment you understood rather than the one that you couldn't explain. This is the situation as it existed I guess about a year ago (at the Vu-graph). And the choice was to do this. This is Sommer, Thomas, and Hipple, Sanders and Turberfield, Boyne and Franken, Mamyrin and Frantsuzov. Well, what have we seen here this week? And, lo and behold, the nice two-humped camel that Lew Branscomb objected to so eloquently at Munich two years ago has completely disappeared. We're right back to choosing the average as it existed roughly two years ago. This has reminded me of a verse that I saw somewhere. (*Laughter*). I'm not sure where it came from. It's not in the approved format; it is a couplet that goes:

> "*The last temptation is the greatest treason*
> "*To do the right thing for the wrong reason.*"

But Cohen and DuMond and Taylor, Parker, and Langenberg I think have a reply to that:

> "*The greatest frustration that man can seize on*
> "*To do the wrong thing for the right reason.*"

(*Laughter.*)

D. N. LANGENBERG: We'll take that even if it is in the wrong format. (*Laughter.*)

B. N. TAYLOR: There is one point about the μ_p/μ_n values which people have generally seemed to overlook. What Don was alluding to was the fact that there was an indirect low value of μ_p/μ_n which came from the Faraday. Thus, there were really three low values, not two. We put our eggs in the low value basket because we felt that since the Faraday value came from a different source and supported the other low values, the low value direction was the way to go. I think the experimental evidence in favor of this direction was a little stronger perhaps than has been indicated so far.

A. H. COOK: I'm putting up my hand now because I can at least raise my status as a literary cognoscente by telling Dr. Cohen that his quotation is from T. S. Eliot, "Murder in the Cathedral," a speech by St. Thomas à Becket. (Laughter.)

But I do have a more serious purpose, and I'd like if I may to make about four points. The first is that it is very important to have a consistent set of data and some consistent way of adjusting experimental results because only too often, not only in the history of physics but in the history of astronomy also, people have produced spurious effects simply because they worked things out with inconsistent input constants. The history of science is scattered with

these sorts of things, and I don't need to go into that.

The second point—and this again is very clear in astronomy where the situation with which physicists are now confronted has confronted astronomers for many years past—is that you want to do a least squares or similar adjustment in order to find out how consistent the data are. And for this reason I am very unhappy about the suggestion that people doing adjustments should alter the error estimates made by the observers themselves. One of the points of these adjustments is to find out whether the experimenters (or the observers in the case of astronomy) were correct in their assessment of the uncertainty. And if you arbitrarily go and alter things so as to make χ^2 look a little better, then you're missing a lot of the point of doing the adjustment. I felt very sympathetic towards the remarks that Professor Franken made. I think the burden of his remarks was that physicists don't really assess their uncertainties correctly and put too large uncertainties on their results in order that they may not be caught out. I'm reminded of a story that Sir Harold Jeffreys told when he lectured on the subject of probability at Cambridge. He had taken what was then a rather well known set of calculations by Sir Arthur Eddington of all the fundamental constants from very few input data, based, you know, on the general theory of quantum electrodynamics as set out by Eddington. Well, Jeffreys claimed that he had compared these calculated values with measured values of the same constants, and he calculated a value of χ^2, and χ^2 came out to be extraordinarily small compared with a number of degrees of freedom of about 10 I think. Then Jeffreys pointed out, what is often forgotten by physicists but not so often I think by biologists, that a very small value of χ^2 can be as suspicious as a very large value, and he said either Eddington has fudged the results—which no one can check because no one knows what Eddington's theory consists of—or (and he thought this was a more polite thing to say) physicists commonly put too large uncertainties on their results. They overestimate their errors. This is really a very serious thing to do because it does help to conceal what is the real point of experimental physics, in this field at least, which as I see it is to track down one's systematic errors. Physics advances by people discovering systematic errors. This is almost what physics is about in the area of fundamental constants. And anything that throws a smoke screen over systematic errors, whether in the procedure of the adjustment or in the way in which people quote their errors, is certainly to be deprecated.

A further point: I doubt if it's necessary to emphasize in this audience what least squares is based on. It's based essentially on a Gaussian distribution of errors and—and this is an extremely important point—the absence of internal correlation between sources of error. People sometimes discuss whether a distribution of residuals is Gaussian or not. I regard this as almost entirely irrelevant in physics where the number of degrees of freedom is small and it's impossible to answer the question statistically—it's often possible in biology but not in physics—whether you have a Gaussian distribution or not. But very frequently one has a situation where you have internal correlation between errors. That is to say, you do not have independent sources of error. They are not systematic. You can't tie them down to some sort of formula. But they prevent you applying—well, I say they prevent you applying least squares—the hypothesis of least squares. And in consequence one's estimates of errors from least-squares are too low very frequently.

And, finally, if I may make a plea to physicists—again biologists almost always do this and astronomers sometimes—it's not sufficient to quote a standard deviation. That is not all the information that the statistician or your colleagues want. They also want to know how many degrees of freedom there are on which you based that because that tells other people how much reliance to put on your estimate of standard deviation.

C. Eisenhart: I'd like to comment on I guess several points of Dr. Cook. The first is that I agree with his point about putting on too large uncertainties in the following sense: I feel that when an experimenter puts down his various components of uncertainty he should separate into different sections (a) the statistical part, (b) the parts that describe the additional variation that would accrue if *he* repeated the experiment "in the large," i.e., after disassembly, cleaning, reassembly with replacement of worn or broken parts by new parts of the same kind, etc., and then over in another place, (c) those particular components that are constant for everybody who uses the same method but would be different if you changed the method. (A pendulum and dropping something are different methods of doing a gravity determination.) The reason for such a separation of components of uncertainty is that when people are comparing results with one another they should not include in evaluating the uncertainty of the difference those components which the results share, and these shared components will not be averaged out when the results they affect are averaged. They are not reduced. They are still there. So I'm glad to be able to agree with Dr. Cook to begin with.

My next point has to do with one of the remarkable achievements of Gauss. In his first development (1809) of the method of least squares he started from the principal arithmetic mean as an axiom, showed that the law of error that made the arithmetic mean come out the most probable value was the so-called Gaussian law of error, and then developed therefrom his formulation of the method of least squares, which added to what Legendre had done by introducing measures of precision, and so on. In the course of this development he became convinced that it was impossible to develop a procedure for finding the "most probable value" of an unknown quantity without hypothesizing a par-

ticular law of error. Then in 1821 Gauss developed a new version of the method of least squares in terms of mean square error and minimum-unbiased linear estimates (i.e., linear functions of observations with minimum variance) which makes no assumption about the laws of error involved other than that they have a finite second moment, and that the individual errors be independent. [For fuller details on Gauss's two very different "proofs" of the method of least squares, see Churchill Eisenhart, "The meaning of 'least' in least squares," Paper 4.5 in NBS Special Publication 300, Volume 1; and Churchill Eisenhart, "Carl Friedrich Gauss," International Encyclopedia of the Social Sciences, Vol. 6, pp. 74–81.]

Later (1826) he extended this development to situations involving correlations that come from constraints on the measurements. A century later A. C. Aiken (1938), and also E. R. Cohen (1953) here, extended the theory and methodology of least squares to include situations where there are correlations between errors of the individual measurements or observations; these extensions of the method of least squares are not critically dependent upon Gaussian distributions of underlying errors. I guess that's enough for the moment.

J. S. Thomsen: I'd like to comment on the point about changing the experimenter's errors. There are two instances in which this arises: One, in starting out with an evaluation—I'm not too familiar with details of the situation that confronted Taylor and co-workers. But back in the '50's in looking at the situation you found some experiments stated in terms of probable error, some with a purely statistical error with no allowance for systematic effects, and, if I recall correctly, one or two saying the error was believed to be several times the probable error with "several" left undefined. And when you have a situation like that, I think that honesty compels you to make some adjustment and make an attempt, however crude, to put them on a common basis before you run it through the mathematical machinery.

E. R. Cohen: I don't think putting them on a common basis is changing the error. That's merely changing from centimeters to inches.

J. S. Thomsen: Wait a minute. When you say several times the probable error—that was one of the experiments we had to contend with. Another one, I think, gave a limit of error without stating what this implied—so you do have to make some value judgments there in deciding what number is equal to "several," whether limit of error is really 3 sigma or what it is.

The other situation is one such as I was just talking about in changing the error assignments on the proton moment—I think perhaps this could alternatively be stated as reweighting, changing the weights—and my contention there was that by reducing the weight of one experiment by a factor of 3 or 4 you did less violence to the experiment than by rejecting it entirely, which is essentially reducing the weight to zero.

L. M. Branscomb: I want to make a comment and ask a question.

First, with respect to Professor Langenberg's question, if the panel didn't answer I wish they'd give an answer. Surely the specialist's version of the recommended constants would not be hidden. It should in fact be published and there should be a lively literature. I agree thoroughly with Barry Taylor and others that have spoken out for the value of the evaluation process in terms of what it tells you about physics. That should be published and there should be good, lively competition in that area if there are people willing to do it. But the question I think is a very good one. Following the philosophy that Peter Franken has advanced, and that has been advanced by others, the experimenter should be giving an honest statement based on the betting theory in effect about his confidence in his own work. The user of a set of recommended constants is in much the same situation. He wants to know the basis for confidence that he should have in these numbers.

And if you take that philosophy, then it seems to me you really cannot in conscience throw away any information where "information" is information of value and not just numbers on a piece of paper.

E. R. Cohen: You have to separate signal from noise.

L. M. Branscomb: Yes. There is a growing interest in a methodology for making decisions in the absence of certainty, and it seems to me there is a real analogy between the formal theory of decision analysis and the approach that Pete has taken, which in effect says there are some human subjective judgments to be made and you have to learn how to quantify some amount of subjective judgment in a mixture with some objective judgment. You can evaluate your ignorance through that theory. That's what we are basically after—a quantitative estimate of ignorance.

I wanted to ask two questions. I wanted to ask a question of the audience and a question of the panel.

The question to the audience is: Do experimental physicists have an excuse that their work may be difficult to evaluate in that the journals don't accept the level of detail and documentation, and so forth, that is necessary to really give the professional evaluator a fair chance?

My question to the panel is: Suppose you adopt Bender's philosophy. What do you do with a paper in which there is simply insufficient information? In that case are you justified in ignoring the measurement because there is simply not enough description in order to evaluate the result?

W. H. Parker: With respect to one of those questions, I can give a partial answer. That is the question of whether the journals allow you to publish the full details.

I think at least from my own experience the answer is yes. The journals will allow you to publish full details. But often there is a financial problem when you try to publish the number of pages that you're

talking about, and I do hope you could do something about that. (*Laughter*)

P. L. Bender: Is John Hipple here by any chance? Someone told me that in the original paper on the omegatron he was prevented from giving more details by the journal saying that the article had to be compressed. I hope the situation isn't still true, but it apparently was at that time.

Would anyone else on the panel like to answer the last questions?

C. Eisenhart: Well, I'd like to answer Dr. Branscomb's question about the paper that doesn't have enough information more or less along the lines of Barry Taylor's strategy in other directions. I would suggest that you pay no attention to it. That might provide some incentive to people to give more details.

J. S. Thomsen: I would say on that that probably most of us who have tried evaluating at least at some stage have consulted various experimenters either through letter or personal visits, and, if the paper is not a complete blank on pertinent information, frequently you can fill in the gaps that you need in order to make a reasonable estimate.

D. N. Langenberg: This is a very useful technique and I think we all use it. However, one often runs into geographical and political complications, as in the important case of Mamyrin and Frantsuzov, with whom it is very difficult to communicate.

J. M. Cameron: I'd like to ask a question of Dr. Branscomb. Would our boss have us put on our certificates of our measurements estimates that were not based on observations, estimates entirely based on our subjective judgments when they could not be validated, or if it's a case of, say, a court hearing of some sort, would you have to take that type of evidence and throw it into hearsay? It used to be that in the computers you had garbage in and garbage out, but now you have a feeling anything in and garbage out. (*Laughter*) If we base all this on judgment without any repeatable evidence, or any sign of how we would repeat it or verify it, are we not leaving ourselves open for some problems?

L. M. Branscomb: You're mixing two different questions. We're right back to what Peter Franken said. In the question of a certificate, which is for a calibration or test, which is really outside the domain of this conference, there may be some legal necessity for having an answer that is "correct." Translated into the domain of this discussion that means a 99.99 percent confidence limit. And nobody on the panel has yet proposed that the fundamental constants can be described usefully to anybody within such limits.

In fact, what is advocated here is that you give the reader the best value you can arrive at by the sum of your intelligence, and a statement of error. Most people propose something that gives them 70 percent confidence range, and I haven't heard anybody advocate also giving out a 99 percent confidence range and a 99.99 percent.

P. Franken: I think there are some times when in fact it is useful to try to establish a limit of error which has the operational definition that it would really be astonishing if the number lay outside these limits, and that, in fact, in some cases when you're trying to contrast experimental results with a theory, it's very useful. There have been a number of occasions in the last couple of decades. But I'd like to respond to the comment just made and also once or twice earlier. Some of you really seem to be afraid of situations in which evaluations are, to use a coy word, subjective—or, to be even more explicit, that involve judgment. I'm astonished.

B. N. Taylor: With regard to the earlier question concerning two sets of constants, as I tried to point out in my introductory ten-minute talk, there is probably no need for two sets. It would probably be perfectly acceptable to start out with one set, no matter how subjectively arrived at (I would advocate expurgating inconsistent data), and chop off extra digits so that it is a limit-of-error set. This would serve the complete needs of most workers, and the original set would at least serve as a starting point for the rest.

J. Geist: Along the lines of subjective evaluations, would you care to comment on the Stefan-Boltzman constant measurements that were done up until around 1929 and petered out when Birge published a value for σ based on measurements of c, h and k that agreed with the experimental measurements of σ by Coblenz first published about 1915. Errors in h and k were discovered later. The estimated uncertainties were about a tenth of a percent. The errors were on the order of a percent. I feel that there was subjective evaluation of uncertainties, particularly after Birge's values agreed with Coblenz's. Don't you think there is the possibility of this kind of problem arising if much emphasis is put on subjective evaluation of errors?

P. Franken: I'm sort of torn between two answers. One of them is: That's a fascinating question. Do you have another one? And the other is: I couldn't fail to disagree with you less. (*Laughter.*)

J. S. Thomsen: I'd like to comment on the last two questions. First, on the radiation constants, I looked at that at the time of our evaluation during the '50's when we tried to check out the adjusted results of all the constants against the more crude earlier experiments, and my only conclusion with the radiation constants was that the situation was a mess. And I think they disagreed with the adjusted results by more than almost any other way of getting at the atomic constants.

I think as far as this subjective phase goes—and I suppose this is a nice word to describe the amount of witchcraft in an experiment—the point is that there is something subjective in all of these numbers we play with in the atomic constants game once we adopt the concept of betting odds with systematic errors. You automatically are talking about something subjective there. And one of the most subjective things is: Are the sources of errors which you have examined and taken into account more or less in estimating the systematic error—e.g., are these

really the dominating systematic errors? If you assume one is negligible or unimportant, this may be the most subjective judgment of all. And simply because you can put numbers on some parts of the systematic errors that you state and you can establish some repeatability, say on the temperature variation or something like that, does not prove that you have taken into account all of the crucial systamatic errors. So it seems to me in most of these experiments that we are kidding ourselves if we think there is any input datum, no matter how numerical it is, which is free from a subjective element, and therefore perhaps we shouldn't be too alarmed if Pete suggests doing some further subjective judgment on the numbers.

P. L. Bender: I'd like to make one other comment at this point on the question of whether one should just make the uncertainties on the generally published values large enough so there is no danger of anyone falling outside. There is a real problem in this respect. I have talked with a number of people in the field of geodesy concerning ways in which one can hope to get higher accuracy than is presently obtained, and the reaction that one gets from a substantial number is that there is no point in improving the accuracy because the speed of light isn't known well enough. And the number that is always quoted for the uncertainty is 1 part in 10^6, which was the 3 sigma value that was quoted in the last adjustment that was published. Now, I think that here is a case where by expanding the error to be safe—it shouldn't work this way but I think it actually does—one is to some extent impeding progress in measurement science.

E. R. Cohen: I think people ought to go back and learn how to read. There was a clear warning, admittedly on the pocket card it's in rather small print—but if you read the full-size 8×10 pieces of paper, I think it's quite legible and understandable. I agree that 3 sigma was really a mistake and for just the fact that apparently people can't read. But the reason for the 3 sigma was perhaps the gnawing problem that we had rejected data and therefore we recognized that there was indeed a possibility that we could be way off base, as indeed we were.

As for the choice of two sets of numbers, I don't think there is any reason for publishing a set for children and a set for adults. People will take as much as they can digest. And they do it without worrying about the numbers that they threw off on the other end of the list. They will still use 1.05 for $\hbar$.

A. H. Cook: I think Pete's example of the question of the geodesists reveals great lack of understanding among them because in fact all geodetic measurements and all measurements of celestial mechanics are made in terms of the velocity of light. It doesn't matter what the velocity of light is. It's the relative measurement with respect to it, the accuracy with which you can make timing measurements either on the ground or in the sky that matters, and not the value that you take for the velocity of light. And that's a thing that people do really need to be very clear about. Where an uncertainty in a basic standard enters and where the measurement procedure is eliminated is most important in the type of work we are considering.

May I just take a moment to come back to systematic errors because I think in an earlier comment I may not have made myself entirely clear. I was talking about correlation between errors. Where one knows what the correlation is formally, then there is, of course, a well established procedure for manipulating the least-squares equations to deal with it. But the situation I have in mind is where correlation reveals itself in a much less formal way through relations between blocks of variances. And this is something where it's difficult to tie down the correlation, but these relations tell you that it's there and they are a warning sign. Systematic errors are something that we ought to try to do away with, of course. One way is to refine our understanding of the experiment so that they appear as corrections—that is to say, something that you can determine. But if you do that, you will still find yourself left with things which are pushed, as it were, the other way and have become much less amorphous and reveal themselves as correlations between different groups of observations which you cannot tie down in a formal way but which nonetheless make the final uncertainty of one's result greater than you would suppose if you would just proceed by assuming that there is no correlation.

L. A. Currie: Regarding the question of correlation and systematic error, I should like to call attention to a basic tradition in chemical analysis where one attempts to apply scientific judgement in advance of the experiment.

According to this tradition, which has proved to be extremely beneficial, the analytical laboratory is required to employ at least two methods of analysis which are believed, for sound chemical reasons, to be uncorrelated. Such an approach provides a reasonably objective means for the detection of systematic error, but it may result in embarrassment of the sort discussed by Dr. Eisenhart in connection with a steel SRM.

D. N. Langenberg: Apropos the three-sigma table, although I am fully in accord with the idea of publishing tables for adults only, I can quote a counter example where that Cohen-DuMond three-sigma table was useful, and that's my own case. The very first time I took a look at a table of constants to see how well you would have to do on a measurement of e/h to have a significant result, it was that three-sigma table. I was not fully aware at the time that it *was* a three-sigma table, and I'm not sure I would have been so encouraged about the prospects of measuring e/h if I had known it was a three-sigma table.

W. H. Parker: From the same experiment I'd like to go just the opposite way. Using the one-sigma table, there is a certain thrill and exhilaration that comes from discovering that your answer is four

sigma away from the accepted value that you would not get if you found that your answer differed by only 1.5 times the assigned error. So it goes both ways.

R. T. Robiscoe: I would like to address a question to the panel. Would you favor a formalized method of communication between the experimenter and the evaluator concerning those experiments which may have a relationship to the fundamental constants? Some such thing as a request perhaps by Taylor, Parker, and Langenberg or whoever will handle these constants that the question of systematic and statistical errors be discussed in somewhat more detail than one gets in a typical publication?

B. N. Taylor: I think that that could be quite prejudicial to any new person who might want to get into the field. It is probably not a good idea to restrict communications to specific channels. People come in and out of this field as time goes on.

P. Franken: But it does cost less per page.

B. N. Taylor: That's probably true. But anyone who gets into the constants-adjusting field is obligated to get in direct touch with the experimenter if more information is needed. The information should be kept available by the experimenter (in notebooks perhaps), so that if it isn't published, it can always be given when needed. That is the number one requirement.

E. R. Cohen: All I would say is let me know that you have an experiment that you think is worthwhile and you'll hear from me with some questions.

J. Stern: I just want to throw in the thought that in a good many situations the resolution between disparate values can come about if you really take some trouble to close the measurement loop. Take the example that Dr. Eisenhart presented about the two methods of measuring nitrogen content. The thing to remember is that you're not interested primarily in nitrogen content. You're interested in properties of steel. And you have these two methods which don't really define the composition with respect to nitrogen but that define something you're measuring in which the apparatus and the method also play a part. What you are really interested in is which of the results, nominally nitrogen content, correlates best with the properties of the steel that you're interested in. I think this is a general sort of thing that also applies to the fundamental constants because they have purposes. You get them for certain uses, and they interact with other constants. And this is what you need to keep in mind.

D. A. Owen: There is perhaps another point that should be brought out. When one uses quantum electrodynamics to determine some fundamental constants, one has to be able to assign errors to the terms which cannot be calculated, and I think that perhaps not enough work has been done to systematically estimate the uncalculated terms. Perhaps theorists should get together and have some agreement on how this should be done.

P. L. Bender: Here is a case in particular where Peter Franken's comment is certainly true that the evaluator cannot make the judgment; that the person doing the calculation certainly has to do it.

E. R. Cohen: There are, of course, two different kinds of terms that are uncalculated. One is the higher order terms in some sort of expansion that you haven't gotten around to yet or you don't have the money to support five graduate students for ten years to carry through. And the other is just the effects that you haven't thought of. This perhaps is the more serious one because there's nothing you can do about it. It's either the theoretical terms that you haven't thought of or the experimental systematic errors that you haven't thought of that are the ones that are really the serious problems in an experiment or in a calculation. Either estimating the higher order terms in an expansion that you know something about or making estimates of systematic errors where you have already thought of what the systematic error might be, these are the simple ones to handle, and I think there's really no problem there.

D. A. Owen: I was referring to the fact that different people estimate the higher order terms in quantum electrodynamics in different ways, and it's very hard for somebody else who comes along and looks at these values to make any comparative judgment unless there was some unanimity or confidence level given for the uncalculated terms.

B. N. Taylor: I would like to point out how difficult that is. If one looks at the (theoretical) upper limits of uncertainty assigned to the theory of the Lamb shift since Bethe's calculation in 1950, one finds that every estimate was incorrect.

D. A. Owen: I believe our positronium paper has a reasonable way of estimating the next order. I'm not saying this should be the way to estimate it, but it tries to make estimates of lower orders and see how they work out, and then uses this scheme to estimate the next uncalculated order. Perhaps if this was done systematically to find out some way of empirically trying to make some estimates it might prove of some value.

E. R. Cohen: Well, on just this point, this is the sort of thing that Drell and Sullivan tried to do, to estimate the effects at each order, then evaluate the order, to get a feeling of how well they could rely on their estimate of the order that they didn't calculate. But now there are additional terms beyond that order which come from different mechanisms, and if you don't know what those additional mechanisms are, the estimates that you make as to what the magnitude of all the uncalculated terms might be don't have too much significance. It's unfortunate. That's the sort of thing you have to live with. And that's one of the reasons for trying to understand the consistency of the constants.

B. N. Taylor: I still think it does not let the theoretician off the hook. I think he is in no different a position than the experimentalist. He has to estimate the uncertainties of his theory just as well as he can, just as any good experimentalist examines the uncertainties in his experiment.

G. W. Erickson: Barry, I think that I must

differ with you here. The theorist *is* in a different position. Generally a theorist will present his calculation of a particular term to a given order (which often is an exact calculation, presumably correct) but his having done *that* calculation does not necessarily put him in a better situation to estimate the *other* terms. Perhaps somebody who hasn't done that particular calculation but knows more about the theory might be in a better position to estimate the uncalculated terms. In contrast, the experimenter who has done the experiment is usually in the best position to estimate what he has left out. I think the theorist is in a very different position.

P. VIGOUREUX: Dr. Cook said, if I understood him correctly, that it would be better if experimentalists tried to express their results in terms of random rather than systematic errors. One has to remember that some errors are random for one person and systematic for another. (*Laughter.*) If in the working out of my result I had to use a result of Dr. Cook's, then I call any possible error of that result a systematic error because it was with me all the time. (*Laughter.*) In a current balance I have to make use of the acceleration due to gravity. And whatever value I take, if there's an error in this, my results are in error all the time. Therefore I call that a systematic error. I think that is correct. It's not possible for one person to perform every single experiment to prepare his own determination of something. One has to take other people's results. So if one uses those results and they are the same all the time, they are systematic to this particular investigation.

P. FRANKEN: They are not systematic. They're nagging. (*Laughter.*)

C. EISENHART: I would like to get in one more remark before the time runs out. I would like to compliment one of the Bureau slides that was shown a few days ago in conjunction with that many-author paper presented by Dr. Witt. The slide I am referring to is the one that showed the observations in chronological order; the top of the slide identified the day on which each observation was made, and what cell was used. Down at the bottom another little legend showed when certain changes in apparatus had been made, and so forth. My feeling is that this is the real way to present your experimental data. With a presentation like this and availability of modern high-speed computers with procedures for analyzing data like OMNITAB [See "OMNITAB II, User's Reference Manual," by David Hogben, Sally T. Peavy, and Ruth N. Varner, NBS Technical Note 552; and "OMNITAB II, Magnetic Tape and Documentation Parcel," by the same authors, National Technical Information Service, U. S. Department of Commerce.] and certain others, the data can be really scrutinized. Such scrutiny often reveals unsuspected correlations, unsuspected tendencies to cluster, and so forth, that are not revealed by just visual examination.

D. N. LANGENBERG: I couldn't agree with that more. I would like to point out on the question of publication, however, that the Physical Review (for one) objects strongly to publication of explicit time data, like days of runs. We ran into that problem on the Parker measurement.

C. PAGE: That is easy. We'll publish it ourselves.

P. L. BENDER: Churchill, I would like to ask one question just for the record. How do you feel about adding systematic errors of different kinds as the sum of the squares?

C. EISENHART: You want to know how I feel about adding systematic errors as the sum of the squares? Well, in the first place one has to think about what you mean by the systematic errors. And in this one should pay attention to Dr. Vigoureux's remark about the error that Dr. Cook contributed to him. Everybody who is going to use Dr. Cook's value is going to have this same systematic error. So I feel that it should be listed as a systematic error in Vigoureux's table and other people's similar tables. My point is that the allowance for this systematic error should not be added in quadrature when combiing results of experiments all of which use the very same value.

On the other hand, when you are combining results from two different methods in which there is a component of systematic error that is fixed for one and either not present at all or present in a different magnitude in the other—the limits that can be placed on the two errors may be the same, but the two errors may be free to take on different values between these limits—then by combining in quadrature you are essentially considering these systematics to be random in the composition. I have no alternative for doing it, and I feel there is a point to doing it.

In section 4.2 of that enormously long paper I wrote on "Realistic evaluation of the precision and accuracy..." [reprinted with corrections as Paper 1.2 in NBS Special Publication 300, Volume 1], I discussed this matter in considerable detail and pointed out that problems arise in converting to a corresponding "standard error" or "probable error," and the choice of rule for effecting the conversion can make quite a difference in the end result. The biggest difference comes—and this is why I would really prefer to use something like a three sigma—if people are estimating the guesstimated kind with what they think is their limit of error, or if they are getting a probable error and then multiplying it up by 2 or 3, they should say what they are doing, but then when you come to the statistical ones, here's where the real bite comes in. The standard error that I got from the Canadian measurements, that squared thing, is an unbiased estimate of σ^2, but if you were to put a confidence limit on it and try to get it out to something that corresponded to three sigma of sigma, so to speak, you would inflate it a great deal, and this would make more explicit in that sum of the squares that Barry Taylor called σ_T^2—it would make more explicit the uncertainty of the σ_m^2.

I'm not trying to dictate to anybody how they should do it, but I feel that by playing with these

things one gets a feeling as to how uncertain the σ_T^2 is.

T. LASHOF: Not being familiar with this field, I was going to keep quiet, but there is one question or one point that I think needs clarification, for me anyhow, and that is that there seems to be a lot of discussion about how to arrive at a weight or an error for each experiment. One possible approach when you have redundant data, which is what you need, of course, when you are going to make a least squares fit, is to omit one item of data, run the analysis, determine the weight for that item or a preliminary weight, and continually go through this until you get a preliminary weight for each experiment, and use these preliminary weights for another iteration of this same process. Eventually you would arrive at a weight (or an error which is equivalent) for each experiment. This eliminates the problem of bias or judgment on the part of the adjuster or misjudgment on the part of the experimenter in presenting his weight values. But it also may lead—and this is where you people come in—it may lead to an absolutely ridiculous set of weights. I was just wondering how this kind of approach might work out.

C. EISENHART: Ted, I've never heard somebody doing it quite the way you have described, but John Tukey proposed (1962) a procedure which he has baptized with one of those names that he's always concocting—"Jackknife"—which is almost like what you described but not quite. What he would do, if I understand it, would be go through and calculate the mean with each one of the individual results left out in turn, say $\bar{x}_{n-1}{}^{i}$, $(i=1, 2, \cdots, n)$, where "i" is the index of the value left out, and then simply take the mean and dispersion of n quantities $n\bar{x}_n-(n-1)\bar{x}_{n-1}{}^{i}$, $(i=1, 2, \cdots, n)$, where $\bar{x}_n$ is the usual arithmetic mean of all n, and treat these as if they were independent results in a t test. And, of course, they are not really independent. But Tukey has shown that this gives a remarkable kind of a confidence limit. It's sort of mysterious, but he's a pretty smart fellow; and there are now quite a few papers on "jackknifing" in the literature of statistical theory and methodology.

B. N. TAYLOR: I would like to make just one quick closing statement. Something Mr. Driscoll said the other day with reference to his γ_p measurements is quite applicable to the entire precision measurement—fundamental constants field, namely, "$\cdots$ I don't know that there is anything sacred about repeating and agreeing within the stated probable error." I think there's a lot to think about in that statement. (*Laughter.*)

P. L. BENDER: I'm sure the members of the panel have enjoyed themselves this afternoon. I'm amazed at the number of people who were here through most of the afternoon and even more amazed at the number that are still here. Thank you very much for coming. I would like to close this session but turn the proceedings over to Dr. Ambler for a last word or two.

E. AMBLER: Well, after a whole week I'm sure that there is not very much left to say. It seems to me a long time ago, some 18 months, when some of us sat around a table and said to ourselves, "Wouldn't it be nice to have a conference of this kind?" I think on behalf of the organizing committee we can say that it has turned out to be everything we hoped that it would be; we hope that you have found it useful also.

The thing that remains to be done, of course, is to get the Proceedings out. So with one final last urging, please let us have the manuscripts; the sooner we get them, the sooner you will get the Proceedings.

I think that is all I have to say, so thank you very much.

CONFERENCE ADJOURNED

Author Index

Conference Registrants

Ambler, E.
Audoin, C.
Bailey, J.
Baird, J. H.
Baird, K. M.
Barger, R. L.
Barnes, J.
Bartky, I. R.
Bay, Z.
Beahn, T. J.
Bearden, J. A.
Bedard, F. D.
Beers, J. S.
Belecki, N.
Bell, G. A.
Bender, P. L.
Bjerregaard, J. H.
Bonse, U.
Bower, V.
Bowman, H. A.
Breitenberger, E.
Brodsky, S. J.
Bruce, C. F.
Burr, A.
Cameron, J. M.
Capptuller, H.
Carlson, E. R.
Chan, Yau
Cohen, R. E.
Cohen, V. W.
Cook, A. H.
Corruccini, R. J.
Crampton, S. B.
Cutkosky, R. D.
Danos, M.
Dean, L.
Denenstein, A.
Deslattes, R. D.
Dicke, R. H.
Douglass, D. H., Jr.
Driscoll, R. L.
Dunfee, B. L.
Eicke, W. G.
Eisenhart, C.
Ensberg, E. S.
Erickson, G. W.
Evenson, K. M.
Faller, J. E.
Field, B. F.
Finnegan, T. F.
Fowler, H. A.
Frank, L. E.
Fuller, E. G.
Fulton, T. A.
Fystrom, D.
Gard, E.
Gevantman, L. H.
Giacomo, P.
Gräff, G.
Grotch, H.
Halford, D.
Hall, J. L.
Hall, W.
Hamer, W. J.
Hammerman. I.
Hammond, J. A.
Hara, Ko
Harris, F. K.
Hart, M.
Hegstrom, R. A.
Hellwig, H.
Helmer, R. G.
Henins, A.
Hermach, F. L.
Hill, J. J.
Hillhouse, D.
Hipple, J. A.
Holt, H. K.
Honti, P.
Horsfield, D. A.
Hughes, W. M.
Huntoon, R. D.
Javan, A.
Johnson, D. P.
Johnson, W. H.
Julie, L.
Kamper, R. A.
Kessler, E., Jr.
Ketting, R.
Khorana, B. M.
Kibble, B. P.
Kincaid, B. M.
Klempt, E.
Kleppner, D.
Knight, D. J. E.
Kose, V.
Kponou, A.
Lamb, W. E., Jr.
Langenberg, D. N.
La Villa, R.
Layer, H. P.
Lea, K. R.
Lee, L. H.
Lee, P.
Leiss, J. E.
Levine, J.
Luther, G. G.
Luxon, J. L.
Madden, R. P.
Maki, A. G.
Masui, T.
McDonald, D. G.
Mielenz, K.
Milne, A. D.
Moore, K. H.
Morrow, J.
Motz, J. W.
Narasimham, M. A.
Olsen, P. T.
Owen, D. A.
Page, C. H.
Parker, H. M.
Parker, W. H.
Peterson, C.
Petersons, O.
Petley, B. W.
Polk, L. F.
Pontius, P. E.
Powell, C.
Preston-Thomas, H.
Ramsey, N. F.
Raybold, R. C.

Repko, W. W.
Rivera, A. T.
Robinson, H. G.
Robiscoe, R. T.
Rossini, F. D.
Roush, M. L.
Rytz, A.
Sakuma, A.
Sauder, W. C.
Saunders, J. B.
Scalapino, D. J.
Schenck, A.
Schweitzer, W. G.
Series, G. W.
Shields, J. Q.
Shugart, H. A.
Shyn, T. W.
Simmonds, M.
Skapars, A.
Smith, S. J.
Sober, D. I.
Steiner, B. W.
Steudel, A.
Stille, U.
Stoner, J. O.
Strain, D. C.
Sze, W.
Tan, D.
Tarcia, M.
Taylor, B. N.
Terrien, J.
Thomsen, J. S.
Toots, J.
Turgel, R. S.
Tyson, J. A.
Uzgiris, E. E.
Van Assche, P. H. M.
Vanier, J.
Vessot, R. F. C.
Vigoureux, P.
Vorburger, T.
Wapstra, A. H.
Ward, J. W.
Wells, J. S.
Wells, T.
Wesley, J. C.
Witt, T. J.
Wu, Y. C.
Young, R. D.
Zamboni, F.

FORM NBS-114A (1-71)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBS-SP-343	2. Gov't Accession No.	3. Recipient's Accession No.

4. TITLE AND SUBTITLE
Precision Measurement and Fundamental Constants. Proceedings of the International Conference held at the National Bureau of Standards, Gaithersburg, Maryland, August 3-7, 1970.

5. Publication Date
August 1971

6. Performing Organization Code

7. AUTHOR(S)
D.N. Langenberg and B.N. Taylor, editors

8. Performing Organization

9. PERFORMING ORGANIZATION NAME AND ADDRESS

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

10. Project/Task/Work Unit No.

11. Contract/Grant No.

12. Sponsoring Organization Name and Address
National Bureau of Standards
Department of Commerce
Washington, D.C. 20234

13. Type of Report & Period Covered
Final

14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES

16. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)

This volume presents the Proceedings of the International Conference on Precision Measurement and Fundamental Constants, held at the National Bureau of Standards in Gaithersburg, Maryland, from August 3 through August 7, 1970. The conference brought together theoretical, experimental, and applied scientists for the purpose of discussing modern techniques of precision physical measurement and their application, along with modern theoretical developments, to the determination of the fundamental constants. The topics covered were: frequency and time standards; length standards; the velocity of light; the Rydberg constant; electrical standards; the proton gyromagnetic ratio; the Faraday constant; atomic masses; the proton magnetic moment; Josephson effects; x-rays; fine and hyperfine structure in simple atoms; lepton g-factor anomalies; the gravitational constants; least squares adjustments of the constants. These proceedings contain the invited tutorial talks as well as the contributed papers presented at the conference. Also included are the post-paper discussions and a panel discussion entitled "Should Least Squares Adjustments of the Fundamental Constants be Abolished?"

17. KEY WORDS (Alphabetical order, separated by semicolons)
Fundamental constants; least squares adjustments; precision measurements.

18. AVAILABILITY STATEMENT

[X] UNLIMITED.

[ ] FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NTIS.

19. SECURITY CLASS (THIS REPORT)
UNCLASSIFIED

21. NO. OF PAGES
543

20. SECURITY CLASS (THIS PAGE)
UNCLASSIFIED

22. Price
$6.00

USCOMM-DC 66244-P71

☆U.S. GOVERNMENT PRINTING OFFICE: 1971 O—416-417