INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International

300 N. Zeeb Road
Ann Arbor, MI 48106

8322346

Galison, Peter Louis

PART I. HOW EXPERIMENTS END: THREE CASE STUDIES ON THE INTERACTION OF EXPERIMENT AND THEORY IN TWENTIETH-CENTURY PHYSICS. PART II. LARGE WEAK ISOSPIN AND THE W MASS. (PART I. HISTORY OF SCIENCE. PART II. PHYSICS)

Harvard University PH.D. 1983

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1983

by

Galison, Peter Louis

All Rights Reserved

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __√__.

1. Glossy photographs or pages __✓__
2. Colored illustrations, paper or print _____
3. Photographs with dark background __✓__
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages __✓__
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) __________ lacking when material received, and not available from school or author.
12. Page(s) __________ seem to be missing in numbering only as text follows.
13. Two pages numbered __________. Text follows.
14. Curling and wrinkled pages _____
15. Other ____

University
Microfilms
International

HARVARD UNIVERSITY

THE GRADUATE SCHOOL OF ARTS AND SCIENCES

THESIS ACCEPTANCE CERTIFICATE

(To be placed in Original Copy)

The undersigned, appointed by the

Division

Department

ad hoc Committee for the Ph.D. degree in Physics and History of Science

have examined a thesis entitled

Volume 1: How Experiments End: Three Case Studies on the Interaction of Experiment and Theory in Twentieth-Century Physics

presented by Peter Louis Galison

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature

Typed name Erwin N. Hiebert

Signature Edward M. Purcell

Typed name Edward M. Purcell

Signature Gerald Holton

Typed name Gerald Holton

Date August 1981

HOW EXPERIMENTS END:

THREE CASE STUDIES ON THE

INTERACTION OF EXPERIMENT AND THEORY

IN TWENTIETH-CENTURY PHYSICS

Part 1. History of Science

A thesis presented

by

Peter Louis Galison

to

The Ad Hoc Committee on Physics and History of Science

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Physics and History of Science

Harvard University

Cambridge, Massachusetts

July 1981

ABSTRACT

How Experiments End:

Three Case Studies on the Interaction of Experiment and

Theory in Modern Physics

This thesis addresses the question: How does evidence from a physical experiment become persuasive to those performing the experiment? Two methods are used to examine case studies in 20th century physics. First, experiments are chosen that at some point involved an error; some of the physicists involved believed a result to be true that we now consider false. This throws into relief their criteria for accepting the original (incorrect) result. Second, in each study two competing groups are analysed. This also helps make clear the standards and nature of what each group considered to be compelling evidence. The three case studies are chosen to represent three very different epochs in 20th century physics experimentation. The first, "Einstein and the Gyromagnetic Experiments: 1915-1925," involves a classical 19th century type of apparatus; the second, "The Discovery of the Muon and the Failed Revolution Against Quantum Electrodynamics," depicts the era of cosmic ray experimentation (especially the 1930's); the third, "How the First Neutral Current Experiments Ended," focusses on several classical high energy physics experiments of the 1970's. A final chapter discusses the changing nature of experimental demonstration in 20th century physics, drawing from the case studies. Revised, published versions of the three case studies may be found respectively in 1) Historical Studies in the Physical Sciences, 12 (1982), pp. 285-323, 2) Centaurus, April (1983), pp. 22-76, and 3) Reviews of Modern Physics, 55 (1983), pp. 477-509.

Peter Louis Galison, Harvard University, August 1981. Degree: PhD in Physics and History of Science. Note: This is part one of a two part thesis. The second part is entitled, "Large Weak Isospin and the W Mass," accepted April 1983.

ACKNOWLEDGEMENTS

I would like to thank Erwin Hiebert, Gerald Holton, Arthur I. Miller, Edmund M. Purcell and Sylvan S. Schweber for all their help and encouragement during my years at Harvard both as an undergraduate and graduate student. At the same time I would like to thank I Bernard Cohen, who though not directly involved in the supervision of my thesis, has encouraged me in all my academic endeavors. To these people I owe a special debt. It has been a great privilege to have worked with them. More than anything, though, I have valued their concern and friendship.

I have also enjoyed and benefitted from conversations on many topics with Thomas S. Kuhn over the last four years.

Without the sympathetic interest of Steven Weinberg, I don't think I could have approached the problems of the history of contemporary physics. His encouragement has meant a great deal to me.

For help on specific parts of the thesis I would like (in addition to the above) to thank the following: Chapter II. A. Pais, Paul Hoffman, Paul Forman, John Heilbron, A. J. Kox, Franklin Portugal, H.-J. Treder, Hendrik Antoon Lorentz, and the staff of the Niels Bohr Library; Chapter III. Hans Bethe, Bruno Rossi, J. C. Street, and the staff of the Bethe archives at Cornell; Chapter IV. F. Everitt, I. Hacking, D. Morellet, C. Prescott, N. Ramsay, J. K. Walker, V. Weisskopf, B. Aubert, C. Baltay, D. Cline, D. C. Cundy, P. Heusse, R. Imlay, T. W. Jones, A. M. Lutz, A. K. Mann, P. Musset, D. H. Perkins, A. Rousset, C. Rubbia, and J. P. Vialle.

During my year at the Department of History and Philosophy of Science at Cambridge University I was fortunate to have benefitted from many discussions with Mary Hesse, Gerd Buchdahl, Nick Jardine, and John Schuster.

I am grateful too for the wonderful hours I have spent talking about various topics related to my thesis with my friends Terry Castle, Richard Garner, Trude Kleinschmidt, Jaklin Kornfilt, John Preskill, Bob Proctor, Donald Reid, Mark Soldate, and Douglas Stone.

It has been an immense help to have had the Graduate Fellowship from the National Science Foundation and the Junior Fellowship of the Society of Fellows at Harvard.

I would like to thank D. Gold, M. Older, and J. Adler for typing the thesis.

Most of all I would like to thank my parents, Max, and William for a million things. I am very glad to dedicate this thesis to them.

TABLE OF CONTENTS

Acknowledgments ii

List of Figures vi

I. INTRODUCTION 1

Notes 14

II. A STUDY OF THEORETICAL PREDISPOSITIONS IN EXPERIMENTAL PHYSICS: EINSTEIN AND THE GYROMAGNETIC EXPERIMENTS: 1915-1925 15

1. Early Tests of Ampère's Hypothesis 17
2. The Einstein and de Haas Experiments 30
3. Einstein's Theoretical Preoccupations 52
4. Barnett: From Terrestrial Magnetism to Einstein's Error 61
5. The Revenge of Experiment: Stewart, Beck and Arvidsson 74
6. Summary and Conclusions: On Theoretical Predispositions 89

Notes 103

III. THE DISCOVERY OF THE MUON AND THE RESURRECTION OF QUANTUM ELECTRODYNAMICS 119

1. Millikan, Cosmic Rays, and the Birth Cry Theory . . . 121
2. New Support for the Birth Cry Theory 129
3. The Collapse of the Birth Cry Theory 148

4. Bethe, Quantum Electrodynamics, and Cosmic Rays . . . 154

5. The Resurrection of Quantum Electrodynamics and the Discovery of the Muon 171

6. Summary and Conclusions 186

7. Epilogue 192

Notes 194

IV. THE DISCOVERY OF WEAK NEUTRAL CURRENTS AND THE TRANSFORMATION OF EVIDENCE 208

1. Introduction 209

2. The CERN Experiment "Gargamelle": From W-Search to Neutral Current Test 217

3. Background and Signal 225

4. The First NAL Experiment 1A 247

5. The Second NAL Experiment 1A 268

6. Conclusion: The End of Experiments 296

Notes 302

V. AFTERWORD: HOW EXPERIMENTS END 315

Bibliography to Chapter II 327

Bibliography to Chapter III 333

Bibliography to Chapter IV 342

LIST OF FIGURES

Note: Figures are numbered sequentially within each chapter. Thus, for example, III.2 will designate the second figure of Chapter III.

II.1 Maxwell's First Experiment 20

II.2 Maxwell's Second Experiment 22

II.3 Mechanical Analogy to Richardson or Einstein-de Haas Effect 29

II.4A and 4B Illustration of Einstein-de Haas Apparatus with Detail 33-4

II.5 Determination of Damping Constant from Resonance Curve . 37

II.6 Resonance Curve from Einstein-de Haas Paper 41

II.7 Data Points from Einstein-de Haas Paper 43

II.8 De Haas' Pulse Circuit 49

II.9 Mechanical Analogy to Barnett Effect 64

II.10 Barnett's 1915 Experiment 69

II.11 Stewart's 1918 Apparatus 76

II.12 Summary of Gyromagnetic g-Factor Results 90

II.13 Publication Year vs. g-Factor 93

III.1 Millikan's Cosmic Ray Absorption Data 133

III.2 Atom Building Processes 135

III.3 Experiment vs. Theory for Millikan's Atom Building . . . 136

III.4 Anderson and Neddermeyer's Energy Loss Data for Cosmic Rays 181

III.5 Street and Stevenson's Apparatus for
Range-Energy Measurements 183

III.6 Stopping Muon 185

III.7 Summary of Discovery of Muon 189

IV.1 Neutral Current Candidate from CERN-Gargamelle 211

IV.2 Neutral Current Candidate from NAL Experiment 1A 213

IV.3 Schematic Diagram of Associated Event 223

IV.4 Number Events vs. Position in Gargamelle 232

IV.5 First Single Electron Event Found at Gargamelle 235

IV.6 R = NC/CC vs. Position 237

IV.7 Summary Sheet from Gargamelle Scanning Meeting 240

IV.8 Figures from Physics Letters 246

IV.9A Design of Experiment 1A 254

IV.9B Typical Charged Current Event Candidate 254

IV.9C Calorimeter Record of Event in 9B 254

IV.10 Wide Angle Muon 263

IV.11 Comparison of First and Second Versions of E1A 270

IV.12 Hadron Punch Through 272

IV.13 "Gold-Plated Event" 275

IV.14 "No Neutral Current" Draft of Physics Review Letter . . 281

IV.15 Letter from E1A to Gargamelle Reporting
No Neutral Currents 285

IV.16 Cline's Concluding Transparency from NAL Talk Dec. 1973. 289

IV.17 Figures from Physics Review Letters 294

CHAPTER I

INTRODUCTION

Einstein once facetiously characterized the difference between theory and experiment this way: no one believes in a theory except its author, whereas everyone relies on an experiment but the physicist who conducted it.[1] Like many of Einstein's quips, this one hides an important distinction, no doubt in large part grounded on his own theoretical and experimental accomplishments.

In this remark Einstein probably was referring to the more freely speculative theories that abound at the frontier of physics where strikingly contradictory ideas are continually put forward. Since these ideas often extend beyond the experimental data, many of them are soon refuted by new experimental results. At the moment they are introduced, however, each theorist has a complex of reasons for advocating his own approach. Some reasons are technical, others more general guiding principles of what he feels physics should "look like." No one can be familiar with every theorist's "hidden" motivations. As Einstein points out, the innovative theorist often remains the lone advocate of his creation.

Here the style of the speculative theorist stands in marked contrast to that of the experimentalist. To some degree this difference stems from practical considerations. Anyone can repeat for himself the steps taken by a theorist. But since experiments are often large, expensive, and difficult (and getting more so) the repeatability of an experiment may be more of an abstract ideal than a reality. If the experimentalist is to retain the confidence of the community, he cannot afford--as the theorist can--to cast many seeds to the wind hoping one will somewhere take root.

Somehow, in the course of an experiment the experimentalist must convince himself that the effect he is looking at is real, and is not an artifact of the apparatus or environment. This is made all the more difficult as his detailed knowledge of the equipment reveals its frailties to him all too clearly. Most of these difficulties will be overcome during the long period of challenges and modifications the experimentalist makes as the work progresses. When he becomes confident that the result will not "go away" he publishes a highly abbreviated account of his reasons for believing the result is real. Necessarily many of the potential sources of error and their relative importance cannot be discussed in detail. Only the experimentalist can know at first hand how much confidence he can place in the many instruments and interactions between instruments that compose a modern physics experiment. Furthermore only the experimentalist knows at close quarters what effects might mimic the one he is seeking. And so everyone believes the experiment but the experimentalist.

This is of course too harsh. Eventually experimentalists do come to have confidence in their results. It is the objective of this analysis to explore, with the help of three case studies, the way in which experimental physicists persuade themselves that they are looking at a real effect and how this process has changed in recent history. Case studies have the disadvantage of passing over some long-term historical trends, but the advantage of bringing out the details of individual experiments. Only in such details can we see the gradual transformation of experimental evidence from a suggestive hint to a rationally ordered argument and thus exhibit the nature of evidence experimentalists find persuasive. By choosing three clusters of experiments with the approximate dates of 1915, 1935, and 1973, I hope to illustrate how the nature of convincing evidence has evolved over the course of the century.

By a "rationally ordered argument" I have in mind the presentation of experiments in textbooks or review articles where the conditions of the experiment are explained, the results presented, and their interpretation given. In general such presentations will be formulated as support for a theory. The task of the historian is often quite different. Instead of trying to reorder the conclusions in the most logically consistent fashion, the historian is more likely to seek the original conceptual framework under which the experiment was designed and the results understood. Even so, many historical accounts of twentieth-century experiments gloss over the ground between the design of the experiment and the interpretation of the result. In part, this may be because the primary attention of historians of twentieth-century physics has been on the development of theory: quantum theory and special relativity. When the focus is on the theorists, it seems that examination of the experimental work recedes into the background where it can be described merely in terms of design and results.[2]

By contrast, a specific goal of these essays is precisely to explore the stage of experimentation between the experimentalist's original intent and his final conclusions. In this often hidden realm, the experimentalist may have more than a working hypothesis but not yet a rational argument in final form. The apparatus is functioning, the data have begun to accumulate, but the results are not yet secure. Little by little an argument emerges from the data and becomes convincing. At some point certain convictions crystallize. We can ask: What kind of evidence persuades the physicist that he is on the right track, that he has a "real" effect under observation? What role do the theoretical presuppositions each experimentalist must bring with him play in the outcome of the experiment? What role is played by past experimental experience? To answer these questions we must look at the period where evidence is validated, that is at the time between design and outcome.

To explore this validation period I will employ a technique that has long been known to historians of theoretical physics, indeed to historians outside history of science, such as historians of art and literature. It is to pick out certain salient features of a work that seems odd and to bring these elements to the fore. One historian of literature described his approach as one of "underlining expressions which struck me as aberrant from the general usage, [which often] ... taken together seemed to offer a certain consistency."[3] In the history of physics, the aberrant sometimes may correspond to a characteristic style of doing physics such as Einstein's use of symmetry arguments in his 1905 papers. More commonly, the aberrant may seem so only in retrospect as we find the need for revision in the physics of the past.

Historians of theoretical physics long ago recognized that any historical reconstruction that ignores what seems to us (in retrospect) as erroneous will be an inadequate account. As Alexander Koyré wrote in 1939,

> What good is it then to spend time on error? Isn't the important thing the final success, the discovery, and not the tortuous paths that one followed and on which one could have gotten lost?... What is important for posterity is in fact the discovery or invention. Nonetheless (at least for the historian-philosopher) it is the dead end, the error...which are sometimes as important as the successes. They can, maybe, be even more important. They are in fact very instructive by permitting us, sometimes, to grasp and to comprehend the secret paths of [the scientist's] thought.[4]

Though Koyré was writing a preface to his work on Galileo, the search for an account of error has remained a primary methodological precept for historians of more recent physics. Thomas Kuhn put it this way:

> In the process of reconstruction the historian should pay particular attention to his subject's apparent errors, not for their own sake, but because they reach far more of the mind at work than do the passages in which a scientist seems to record a result or an argument that modern science still retains.[5]

If I draw particular attention to some significant false roads taken in experimental physics, it will be partly for the same reasons it is of interest to study them in the history of theoretical physics: they reveal the orientation and "secret paths" of the conceptual development of experimental physics. The point of calling certain beliefs "mistaken," is not to judge past theories anachronistically. On the contrary, one identifies problems and solutions no longer considered significant or correct precisely to understand these forgotten frames of mind in their own terms.

But there is an additional reason for studying mistakes in experiments. Every experiment must, explicitly or implicitly, end with a ceteris paribus clause--the validity of the results is assured only if all other things are taken to be equal. But even in principle no one could check the infinite set of possible interfering effects. Therefore the decision to end an experiment must be taken in some other way. It is of considerable interest to discover how experimentalists decide to end a particular project, and how the criteria by which they do so have changed over time.

When looking at experiments that proceeded as expected, or were subsequently confirmed elsewhere, it is easy in hindsight to remain unaware of the importance of the decision to end the experiment. Later that decision may appear as the inevitable recognition by the experimentalist of something present in nature. However, when an important error is made--one involving physical theories considered to be fundamental--it becomes much more apparent that the decision was not inevitable. Instead, it was actually made as the result of some perhaps very complex set of beliefs, commitments or social pressures. This is the second reason for studying mistakes in experimental physics: they can be used as a tool to uncover the criteria by which the decision is made that a given result is real and that the experiment is over.

The decision, rightly or wrongly, to end an experiment takes place in every experiment but carries particular interest when the issues touch on problems of fundamental importance for the conceptual development of physics. In the first case study, the discovery of the Einstein-de Haas Effect, the nature of magnetism and the structure of the atom were at stake. Einstein and his collaborator, W. J. de Haas, sought to test Ampère's hypothesis that permanent magnetism was due to the mutual orientation of small current loops inside the magnet. With the help of electron theory, Einstein and de Haas put Ampère's conjecture in a sharper form: ferromagnetism was due to the mutual orientation of many electron orbits around their respective nuclei. More significantly, using the electron hypothesis, they showed that the ratio of angular momentum to magnetic moment was independent not only of the number of orbiting electrons, but also of their orbital radii and speeds. A definite value for the gyromagnetic ratio (angular momentum / magnetic moment) could therefore be predicted.

To test this model they suspended an unmagnetized iron cylinder by a fiber into a solenoid. If the current in the solenoid was turned on, the magnetic moments associated with the current loops would be oriented as would their angular moments. Since angular momentum is conserved, they reasoned that the bar would begin to rotate to counteract the increase in angular moment associated with the orientation of the electron orbits.

When Einstein and de Haas performed the experiment (which they did with several different types of apparatus) they found the torque on the rod that they expected. According to quantum mechanics we now attribute ferromagnetism to the orientation of electron spins which is associated with half the gyromagnetic ratio Einstein and de Haas expected from electron orbits. How did Einstein and de Haas found their predicted but erroneous ratio? Independently, at approximately the same time as Einstein and de Haas, an American physicist, S. J. Barnett, performed a series of experiments examining the converse effect: magnetization by rotation instead of rotation by magnetization. When he reduced his data in 1914 his results were equivalent to a gyromagnetic ratio near the prediction made later by quantum mechanics. However, after Einstein published his paper with de Haas, Barnett too obtained data confirming the electron-orbit hypothesis.

In the 1920's, it became clear that the gyromagnetic ratio did not have the value Einstein, de Haas, and Barnett claimed to have found. Only with the advent of quantum mechanics in the late 1920's were the new results explained in a satisfactory way. The first essay addresses the question of why Einstein believed so strongly in his prediction, and how this theoretical belief influenced his, de Haas', and Barnett's interpretation of data and treatment of errors. In short, we can ask how theoretical predispositions influenced Einstein's and others' evaluation of evidence.

The second case study is about the series of cosmic ray experiments leading up to the discovery of the positron and muon in the late 1920's and 1930's. During this time quantum mechanics had been developed but not universally accepted. On one level these experiments illustrate the transition from a classical to a quantum mechanical treatment of fundamental processes as the experimentalists sought to analyze the constituents of cosmic radiation. R. A. Millikan, and many of his colleagues, associates, and students did not approve of the quantum mechanical approach for several years. By contrast other physicists working in the field of cosmic ray physics made frequent use of the new ideas at least in the planning of their experiments. We can compare two groups (broadly speaking, Millikan, S. Neddermeyer, H. V. Neher, I. S. Bowen, H. Cameron on one hand, and H. Bethe, B. Rossi, J. C. Street, E. C. Stevenson, and T. Johnson on the other).

Millikan and his associates denied the existence of several phenomena we now regard as well established. For example, Millikan's group argued that no latitude effect existed, and that no very penetrating charged particles were among the cosmic rays. On the other hand, they found effects that today we reject as nonexistent. Millikan and his colleagues claimed, for example, that the energy of cosmic ray particles came only at discrete values. Only by understanding how these diverse errors offer "a certain consistency" can we identify the theoretical predispositions under which the experiments were conducted. As in the Einstein-de Haas case we can try to reconstruct how strong theoretical predispositions affected the kind of data the experimentalists found persuasive.

The theoretical orientation of Millikan's group can be contrasted with the work of Street, Rossi, Johnson et al. in another way as well. Not only did theory influence the conduct of experiments, the conduct of the experiments shaped the kind of arguments that each group considered convincing. For instance, Anderson and Neddermeyer, from early 1932, began to use individual photographs of energy loss to document their claims about the nature of cosmic rays. Such studies characterized their work for the next five years. Finally, in 1937 photographs of the energy loss of individual particles persuaded them of the existence of a new particle of mass intermediate between the electron and the proton.

Virtually at the same time Street and Stevenson came to the same conclusion: a previously unknown type of particle was present in the cosmic radiation. Unlike Anderson, in his early work Street had not made use of the cloud chamber. Instead, following Rossi's lead, Street had used and improved coincidence and anti-coincidence counters to study the passage of charged particles through matter. Consequently, from an early point in his work Street studied the statistical rather than the individual passage of cosmic rays through matter.

Over the course of several years Anderson's and Street's groups perfected their techniques. Like artisans they became comfortable and adept with certain tools and somewhat mistrustful of others. Anderson and Neddermeyer once remarked of their own preliminary data that they were merely statistical and not as persuasive as clear individual photographs. In fact with Millikan they published an article condemning the use of statistical electronic counters. Street reversed their argument. He later remarked that individual photographs never seemed to him very persuasive. Even his own well-known cloud chamber photographs that led to the first determination of the muon mass seemed to Street less than completely reliable. As he later commented, "anything can happen once."

The contrast between statistical and what we might call "exemplary" forms of experimental demonstration became increasingly evident as particle physics has developed. Detectors like cloud and bubble chambers presented detailed examples of interactions whereas electronic detectors like spark and drift chambers offered higher statistics. The contrasting experimental styles implicit in the two approaches stand out clearly in chapter IV where the focus is on a series of high energy particle physics experiments in the 1970's. Once again we examine two groups that made near simultaneous experimental discoveries. One group was an international bubble chamber collaboration at CERN (Centre Europeen de Recherche Nucleaire at Geneva) and the other an inter-university spark chamber group at NAL (National Accelerator Laboratory in Batavia, Illinois).

We can follow the evolution of the two projects from a search for the intermediate vector boson in the early 1960's to a test of the parton model in the late 1960's. Finally after 1971 the experiments became a test of the existence of neutral currents predicted by Weinberg's and Salam's unification of the weak and electromagnetic forces.

At the same time the neutral current case provides an opportunity to investigate changes in the conduct of experiments that are associated with their increased size and complexity. We may explore how two (for the time) relatively large collaborative groups of laboratories were persuaded that the effect they were studying was real. As in the two earlier cases, the goal of the study is to depict the process by which experimental evidence becomes convincing. Here too, the method is to examine in detail how an important mistake occurred in one of the collaborations. For a period of several weeks the EIA group was persuaded that their experiment provided no evidence for neutral currents at the level predicted by Weinberg and Salam. This period of persistance of the error can be studied to shed light on the theoretical predispositions and experimental practices that were characteristic of various participants in the experiments. Lastly, we can see in both collaborations how the partial evidence that persuaded members of the collaboration was slowly transformed into a publishable form that was endorsed by the group as a whole.

In sum, each of these three cases deals with an experimental issue decisive for some aspect of physical theory: each includes a pair of research groups whose approaches to the problem may be compared and contrasted, and each involved an important mistake that can be used to exhibit the theoretical and experimental orientation of the participants. The sequence of these cases illustrates at least schematically some of the enormous changes experimental physics has undergone since 1915. In the last chapter, "The End of Experiments," I draw some general observations on the character of experimentation as it has evolved over the decades from the 1910's to the 1970's. For as the level of complexity of the experiment has increased, the standards of convincing evidence have also shifted. My hope is that by bringing these three studies together, they will point the way towards the nature of some of these changes.

NOTES

[1]As reported by Dr. Herman F. Mark in an interview with Abraham Rabinovich in "Questions of Relativity," Jerusalem Post, 22 March 1979, p. 7. The remark was said to have been made in a laboratory of Berlin University in 1922.

[2]Notable exceptions to this, for instance, are the accounts by Gerald Holton of Millikan's Oil Drop Experiments, "Subelectrons, Presuppositions, and the Millikan-Ehrenhaft Dispute," in The Scientific Imagination (Cambridge: Cambridge University Press, 1978), pp. 25-83, and the one by Roger Steuwer of the Compton Effect, The Compton Effect (New York: Science History Publications, 1975).

[3]Leo Spitzer, Linguistics and Literary History (Princeton University Press: Princeton, 1948), p. 11.

[4]A. Koyré, Études Galiléenes (Paris: Hermann, 1939), p. 77.

[5]Thomas S. Kuhn, "The History of Science," reprinted in The Essential Tension (Chicago: University of Chicago Press, 1977).

CHAPTER II

A STUDY OF THEORETICAL PREDISPOSITIONS IN EXPERIMENTAL PHYSICS: EINSTEIN AND THE GYROMAGNETIC EXPERIMENTS: 1915-1925

In the midst of his work on the General Theory of Relativity, Einstein became deeply involved in a problem of experimental physics. Collaborating with W. J. de Haas, Einstein helped design and execute a series of experiments to investigate Ampère's celebrated hypothesis that magnetism arises from a current circulating about atoms of magnetic substances. In several careful experiments Einstein and de Haas arrived at just the answer they expected, a result now considered to be almost half of what they "should" have found. Two questions come to mind: why was Einstein so intrigued with this particular experiment? How did it come to pass that the two experimentalists found what they were looking for?

In retrospect it is striking that soon after Einstein and de Haas published their "mistaken" result it was confirmed by an American physicist, S. J. Barnett. Soon the experimental problem was taken up by other experimentalists including G. Arvidsson, J. Q. Stewart, and E. Beck. Still, it took several years before these physicists were able to persuade themselves that the original results of Einstein and de Haas were not correct despite the striking coincidence of the theoretical prediction with the experimental evidence.

1. Early Tests of Ampère's Hypothesis

Ampère's hypothesis, at least in its early form, dates from the very first weeks of the history of electromagnetism. Until Ampère and Oersted demonstrated the magnetic effects of electrical currents, the science of magnetism was principally the mathematical study of the effects of magnetic poles in isolation and coupled together. Even after the initial work by Oersted, a theory as sophisticated as that of Biot and Savart left electrical currents and magnetism as distinct phenomena with different causes--even if currents and magnets were capable of interacting. In 1820, Ampère argued for the unification of the causes of electromagnetism and bulk magnetism: When a magnet and an electric current interact, they tend to orient themselves so as to be perpendicular. Ampère continued,

> Consider now the interaction of an electric current and a magnet and that of two magnets. It will be seen that both come under the same law governing the interaction of two currents, if it is assumed that the current is established at each point of a line drawn on the surface of the magnet from one pole to the other in planes perpendicular to the axis of the magnet. It hardly seems possible to me, from a simple consideration of the facts, to doubt that there really are such currents about the axis of magnets, or rather that magnetization is nothing other than the operation by which particles of steel are endowed with the property

to produce...the same electromotive actions as in the voltaic battery....[1]

Thus Ampère broke cleanly with the widespread belief that a magnet owes its properties to separate north and south pole molecules that are connected in some unknown way. The poles had no special significance other than their position relative to the currents that composed the magnet.[2] Ampère's hypothesis was followed by six years of experimental and mathematical progress culminating in his "Mémoire sur la théorie mathématique des phénomènes électrodynamique." In the intervening years Ampère had reworked his hypothesis about the currents that circulated within the magnet; by 1825 he took them to be molecular in origin. But even more importantly, Ampère had worked out a detailed quantitative treatment of magnetism. Although he conceded that all his experimental predictions could be reproduced by a law (Biot-Savart) based on the two-pole idea, Ampère insisted that only his theory took care of the three interactions: current-current, current-magnet, and magnet-magnet, and related them to one cause. "Those periods of history," Ampère wrote, "when phenomena previously thought to be due to totally diverse causes have been reduced to a single principle, were almost always accompanied by the discovery of many new facts, because a new approach in the conception of causes suggests a multitude of new experiments to try and explanations to verify."[3]

James Clerk Maxwell took Ampère's hypothesis seriously.[4] But to design an experimental test, he needed to know what a current was, and here very little had been established to his satisfaction. Despite the many similarities between the electric current and a flow of a material fluid, Maxwell cautioned that "...we must carefully avoid making any assumption not warranted by experimental evidence, and there is, as yet, no experimental evidence to show whether the electric current is really a current of a material substance, or a double current [positive and negative] or whether its velocity is great or small in feet per second."[5] Still, the possibility that current did involve a material transfer led Maxwell to develop three experiments to exhibit the inertial effects of currents, if such effects existed.

The first experiment Maxwell described in the Treatise dates from at least 1870, when he queried John William Strutt (Lord Rayleigh), "Have you tried whether the sudden starting or stopping of a current in a coil has any least effect in turning the coil in its own plane as it would be turned if the current were water in a tub?"[6] In Figure 1, adapted from Maxwell's illustration, a coil is suspended as freely as possible. If an electrical current involves the transportation of inertial mass, then starting a current through the circuit should cause a change in the angular momentum of the wire. The wire thus should rotate oppositely to the motion of the electricity to conserve angular momentum. It does not appear that Maxwell attempted to perform this test.

Maxwell built the apparatus for the second experiment in 1861 to measure the inertial effects of a constant current.[7] A current is applied across a coil, A. (See Figure 2.) The coil can rotate freely on two pins, B and B'. In addition, the entire armature, D, can be rotated in the horizontal plane as it is attached only at two vertical pins, E and F. The cord visible on the pulley just above the bottom

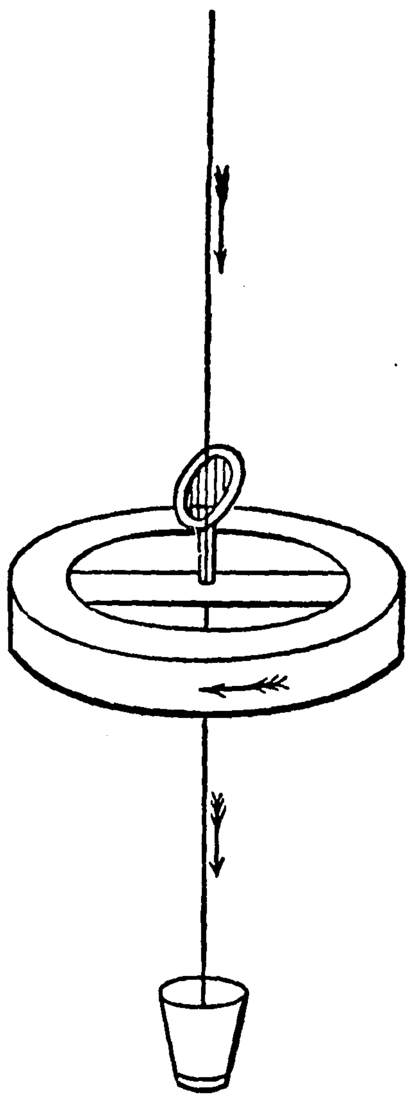
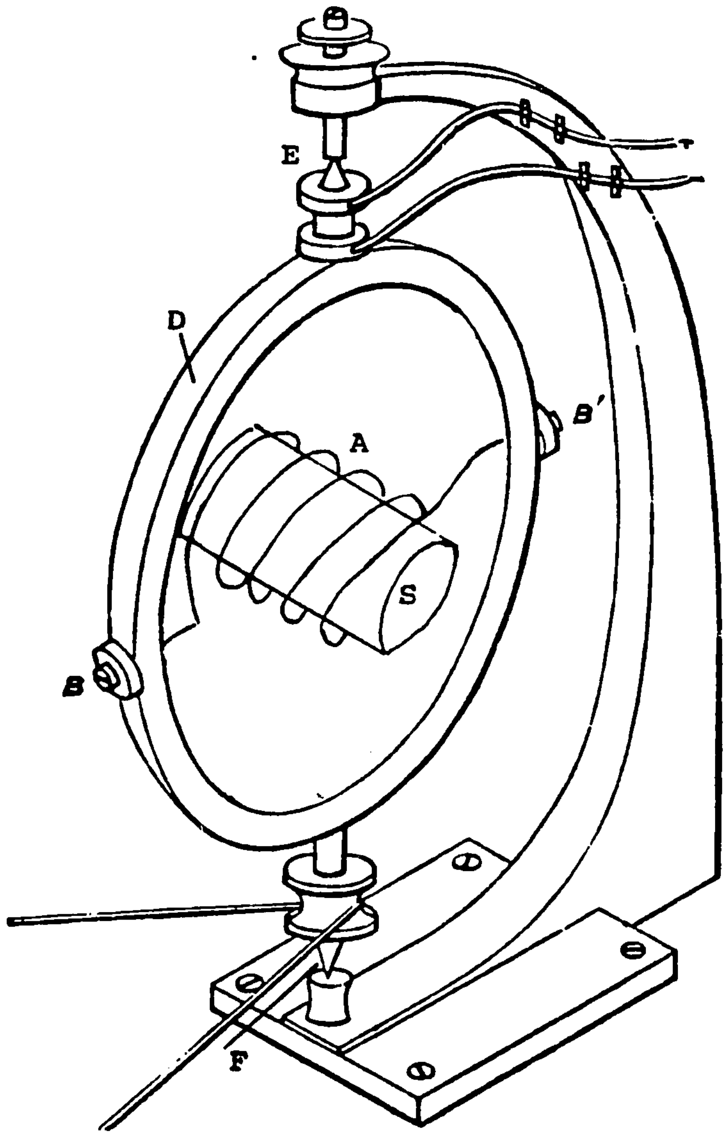


Figure 1. Maxwell's first experiment. Illustration from Maxwell, A Treatise on Electricity and Magnetism (Oxford: Clarendon Press, 1881), p. 201.

Figure 2. Maxwell's second experiment: adapted from Maxwell, *Treatise*, p. 203.


E
D
A
B′
S
B
F

vertical pin, F, is used to spin the armature at a fixed speed. A constant current is applied through the coil by way of two brushes located at the top vertical pin E.

If the current carries momentum, the coil will act like a gyroscope precessing about the vertical axis. Depending on the relative directions of the angular momentum of the gyroscope and the rotating armature, the gyroscope either would tip up or down. The effect is familiar to anyone who has held a gyroscope with its angular momentum vector in a horizontal circle, first in one direction and then in the other. Had Maxwell's experiment yielded a positive result, it would have shown the direction as well as the existence of a material current.

By inserting an iron bar, S, into the coil, A, Maxwell could use his 1861 apparatus to test Ampère's hypothesis: the current through the wire "should" have magnetized the iron bar, orienting the microscopic currents surrounding each "magnetic molecule." This would amplify the tilting effect Maxwell sought to measure.

Maxwell found nothing and explained his failure to do so as follows:

> The chief difficulty in the experiments arose from the disturbing action of the earth's magnetic force which caused the electromagnet to act like a dip needle [a vertical compass]. The results obtained were on this account very rough, but no evidence of any change in θ [the angle the coil made with the horizontal] could be obtained even when an iron core

> was inserted in the coil, so as to make it a powerful electromagnet.[8]

In retrospect we can see that Maxwell had little chance of observing the tilt he was looking for. For only once the currents in the wire and in the magnet were known to be due to electrons was it possible to calculate the magnitude of the effect. Without such an estimate, Maxwell could have no real way to say when he had established a negative result. In 1915 de Haas and his wife, G. L. de Haas-Lorentz, showed that the angle of inclination which would be expected in a device like Maxwell's had a tangent of about 0.00013.[9]

Maxwell's third experiment, like the other two, was designed to test whether the carriers of current also transported inertial mass. A short-circuited coil was given an angular acceleration in its own plane. If the unknown carrier of current had inertial mass, it should lag behind the accelerated coil. The resulting current relative to the coil should produce a magnetic field which presumably could be measured. No indication is given of any actual experiments. Maxwell's idea for this experiment may stem from an apparatus he built and used in 1863[10] to measure the resistance of a wire without reference to the resistance of another sample. In that work, Maxwell rotated a short-circuited wire in the Earth's magnetic field and detected the magnetic field resulting from the convection current. Using a sensitive galvanometer, Maxwell felt he could achieve an accuracy to one part in ten thousand.[11] This may have encouraged him to comment in the Treatise that the null results he had obtained in his experiments on the inertia of current were probably significant:

> ...few scientific observations can be made with greater precision than that which determines the existence or non-existence of a current by means of a galvanometer....If, therefore, any currents could be produced in this way [by accelerating a coil] they would be detected, even if they were very feeble."[12] ...Since, however, no evidence has yet been obtained of such terms, I shall now proceed as if they do not exist, or at least that they produce no sensible effect, an assumption which will considerably simplify our dynamical theory.[13]

After Maxwell, Oliver Heaviside, J. J. Thompson, Joseph Larmor, and J. H. Poynting (among others) continued to pursue Maxwell's interest in the connection between current and momentum. At least before 1897, however, they did not consider the momentum associated with currents to be the result of a transfer of ponderable charged matter. Instead, they ascribed energy and momentum to the electric and magnetic fields associated with the currents. Under this widespread assumption, experiments like Maxwell's on the inertia of currents would not have seemed to the point.[14]

H. A. Lorentz's electron theory broke with tradition by postulating charged, ponderable matter on one hand and the ether on the other.[15] By dividing the two, Lorentz made his charged electrons subject to the same forces as uncharged matter as well as to the forces associated with electric and magnetic fields. One consequence of

Lorentz's separation of charge and ether was that electrons in motion would constitute a material current of the type Maxwell had thought might exist.

In 1907 O. W. Richardson of Princeton University decided to reexamine experimentally Ampère's hypothesis in light of Lorentz's new views on the nature of electric currents.[16] If Ampère's current whirls were simply electrons in orbit around atoms, Richardson could derive some simple relations between their angular momentum and magnetic moment.

Richardson argued as follows: An electron in a circular orbit has angular moment $L = r \times p = r\,(m\omega r) = 2ma$ where m is the mass of the electron, ω is the angular velocity, r is the radius of the orbit, and a is the area swept out per unit time. The accompanying magnetic moment would be, as was known from elementary electrodynamics, simply the electron charge e times the area swept out per unit time: $M = ea$. Therefore, the gyromagnetic ratio--the ratio of angular momentum to magnetic moment--is independent of the angular velocity and the radius of the orbit:

$$\frac{L}{M} = \frac{2ma}{ea} = 2m/e = \lambda.$$

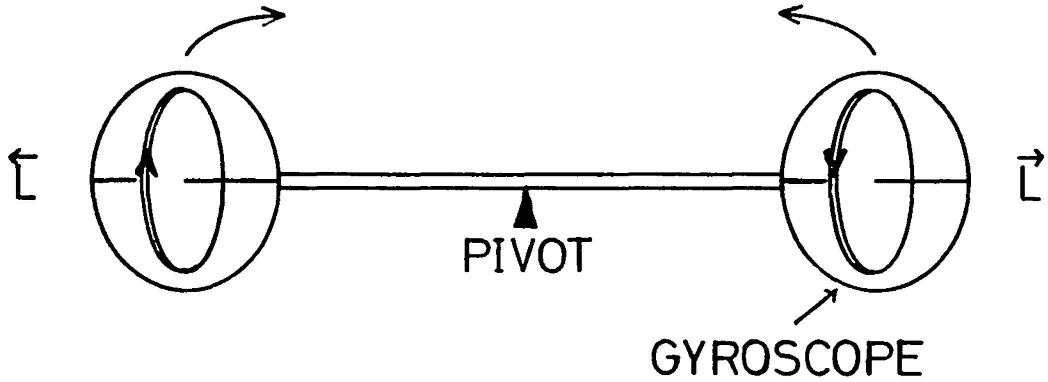
Richardson's formula is easily extended to a general closed orbit. This quantity, L/M, is called the gyromagnetic ratio and in the literature to be discussed is often written as λ. More generally, Richardson calculated this same ratio supposing that both electrons and positive particles were in orbit with different areal velocities.

Richardson, however, did not pursue the more general expression because, "The usual form of the electron theory of matter assumes that the negative electrons alone are in motion and most of the experimental facts seem to be in favor of this conclusion."[17]

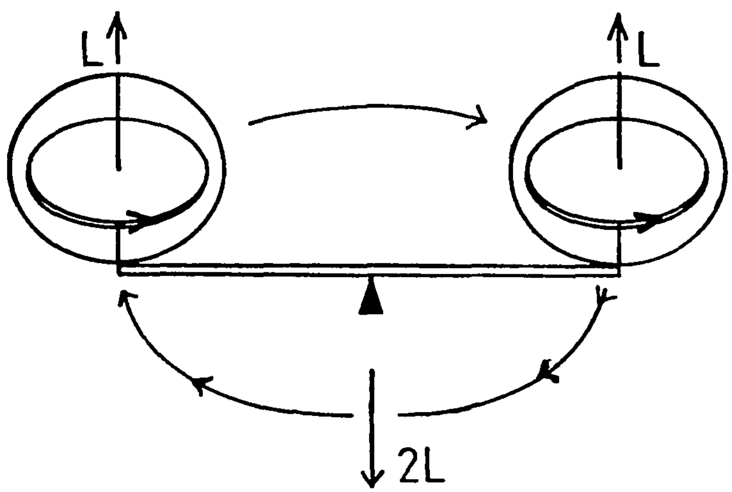
The existence of this constant ratio for all orbiting electrons suggested a simple experiment on which Richardson had begun work. If one could suddenly magnetize a suspended bar of iron, a corresponding change in angular momentum equal to $(2m/e)M$ should set the bar in motion.[18] Like Maxwell, Richardson ascribed his failure to "disturbing effects," without providing details. Nonetheless, as late as 1914 Richardson continued to believe the experiment would yield positive results in spite of the null outcome of his first attempts.[19]

There is a natural mechanical analogy to this effect that may make it clearer. (See Figure 3.) Suppose two identical gyroscopes are spun in opposite directions with the same angular speed and placed, facing away from one another at the ends of a bar. The total angular momentum of the system is therefore zero. Now suppose the bar is placed on a fulcrum around which it can pivot. If the gyroscopes were turned so they stood on end (part 2 of Figure 3), by an agency internal to the rotating arm, the total angular momentum would now be not zero but 2L. To compensate for the change in angular momentum, the whole system would begin to rotate. Analogously, Richardson hoped to orient the microscopic gyroscopes constituted by the orbiting electrons and by so doing cause the macroscopic rotation of the magnetized sample. He found no such effect, however.[20]

Figure 3. Mechanical analogy to "Richardson" or "Einstein-de Haas effect."



$\vec{L}$
PIVOT
$\vec{L}$
GYROSCOPE



L
L
2L

2. The Einstein and de Haas Experiments

Einstein's interest in testing Ampère's hypothesis goes back at least to the period between 1905 and 1909.[21] During those years he regularly met with several other young men (Dr. Hans Flükiger and Dr. Hans Rothenbühler) also interested in problems of experimental physics. Occasionally they met to perform some experiments in the physics room at the Städtische Gymnasium in Bern. Among other projects, they tried to test Ampère's hypothesis experimentally in what seems to have been a rough forerunner of Einstein's later experiment.[22]

Thus when Einstein came to Berlin in April 1914[23] as a full-time member of the Akademie der Wissenschaft, he took the opportunity to pursue his old interest with the much more sophisticated facilities of the Physikalisch-Technische Reichsanstalt in Berlin-Charlottenberg.[24] Einstein's personal friendship with Lorentz and his strong bonds to the Leiden physics community[25] may have contributed to Einstein's choice of Wander Johannes de Haas (Lorentz's son-in-law) as a collaborator in the experimental project. For de Haas, who had arrived at the Reichsanstalt on 1 January 1914 as a scientific assistant (wissenschaftlicher Hilfsarbeiter),[26] the collaboration with Einstein began a long experimental career involving many projects related to this early endeavor.

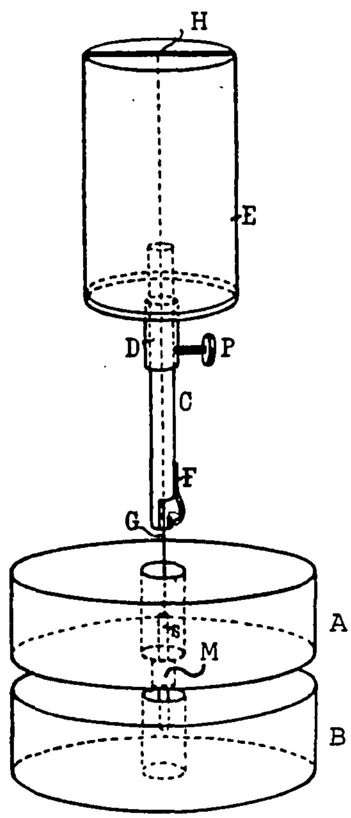
Einstein took an active role in the experimental discussions of the Reichsanstalt during the year. The first report on his own experiment was made in a lecture given by Einstein on 19 Feburary 1915 to the German Physical Society, where for the first time they presented both qualitative and quantitative evidence that Ampère's almost century-old hypothesis was correct.[27]

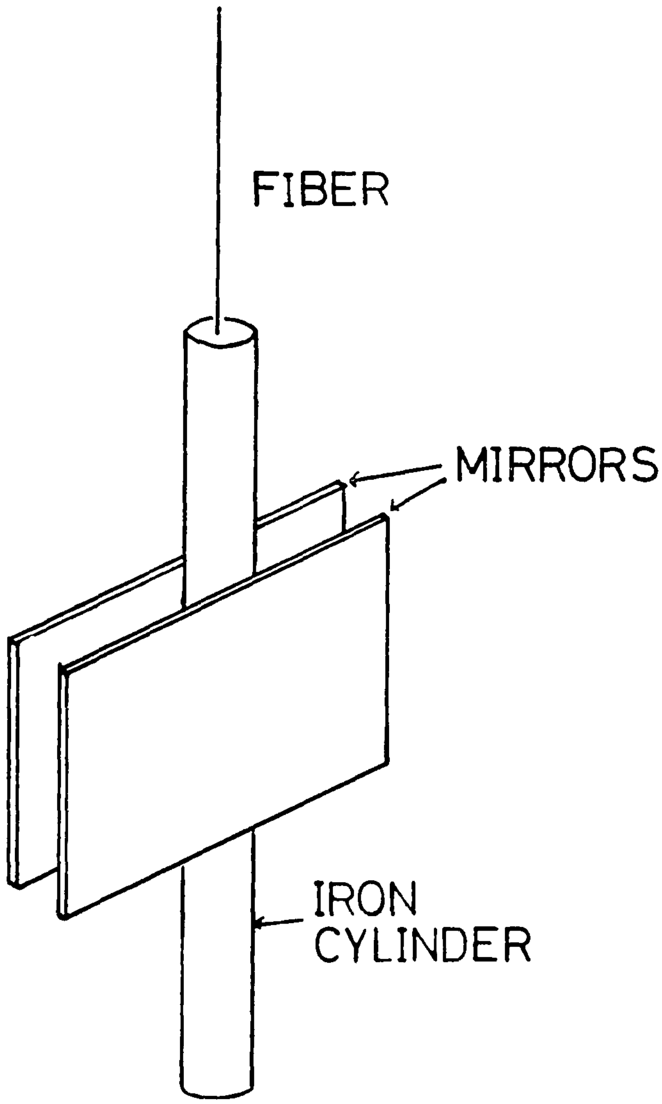
To obtain a qualitative confirmation of Ampère's hypothesis Einstein and de Haas needed only show that by magnetizing a suspended iron rod they could set it in rotation. Unknown to them, the apparatus they were using was based on the same principle as Richardson's. Their chief improvement was to oscillate the magnetic field at the resonant frequency of the bar to amplify the effect. However, since Einstein (like Richardson) also wanted to test whether electrons were responsible for the Ampèrean currents, a quantitative measure was needed as well. It was here that Einstein's theoretical analysis of the experiment gave them the tools to go beyond the simpler experiments of Richardson and Maxwell.

In the first Einstein-de Haas experiment, a fiber G is attached on one end to a crossbar, H, and on the other end to a thin iron cylinder, S. (See Figure 4A.) Two small mirrors, M, are mounted parallel to one another on opposite sides of the center of this iron bar. (See the detail Figure 4B.) Coils A and B surround the suspended iron cylinder, above and below the mirrors, leaving the mirrors exposed to reflect a beam of light from an outside source. The adjustable clamp P is used to vary the effective length of the fiber in order to adjust the natural frequency of the iron bar when in free torsional oscillation.

When an oscillating magnetic field is applied by solenoids A and B, the iron cylinder will oscillate and reflect a beam of light to a screen. The maximum deflection α of this reflected light beam can be measured even if the movement of the cylinder is very slight.

Figure 4A and 4B. Illustration of Einstein-de Haas experiment with detail of iron sample. (Figure 4A from Einstein and de Haas, "Experimenteller Nachweis" (ref. 27), p. 160.


H
E
D
P
C
F
G
S
A
M
B


FIBER
MIRRORS
IRON
CYLINDER

Theoretically, α should be proportional to the torque caused by a change in the bar's magnetization, and inversely proportional to the damping constant. Thus we can write,

$$\lambda = \frac{(\text{constant})\lambda M}{P}$$

where α is the deflection of the light beam, λ is the ratio (angular momentum/magnetization), M is the saturation magnetization of the iron cylinder, and P is the damping coefficient.

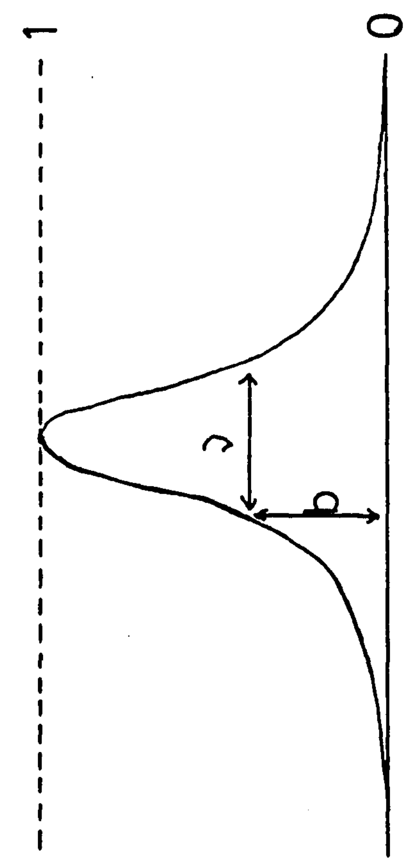
Since α can be measured, and M can be either calculated or measured, only P remains to be determined in order to find the gyromagnetic ratio. In principle P could be directly measured by observing the deflections of successive free swings. Because very small deflections occur this is quite difficult. Instead, Einstein and de Haas chose to measure α when the magnetic field is oscillated at off-resonance frequencies, that is, they measured the "Q" of the system. (See Figure 5.)

Solving the damped harmonic oscillator equation leads in a standard way to an expression for P in terms of the moment of inertia of the iron cylinder, I; the fraction of maximum excursion, b; and ν, the width of the resonance curve for a given value of b.

$$P = (\text{cte.})\, I \quad \nu \sqrt{\frac{b^2}{1 - b^2}}$$

This may be understood in the following way. For a given resonance curve, the greater the moment of inertia, the greater the damping

Figure 5. Determination of damping constant by measurement of resonance curve. For a given cylinder, a narrow spike indicates a small damping constant. ν is the width of the resonance curve at a given frequency and b is the fraction of the maximum excursion of the light beam.


1
0
c
b

constant must be to achieve the same excursion length on the swings. Therefore P is proportional to I. Qualitatively, for a more sharply peaked resonance curve, ν will be smaller at any given height and therefore P will be smaller. This can be understood as follows. If no damping occurred, there would be no way for the off-resonance magnetic field oscillations to couple to the mechanical oscillations of the iron cylinder. However at resonance energy would constantly be pumped into the system and no equilibrium excursion length would be reached. The resonance curve would approach an infinite spike at the resonant frequency and be zero everywhere else. Conversely, as P gets very large, the curve will spread out.

The key measured quantity in the experiment is therefore a good determination of the deflection of the light beam at different frequencies. If it is correct to assume that the damping term is independent of the excursion length, the calculated damping coefficient P should remain constant. Frequency measurement, however, was not nearly so routine a matter in 1914 as it is now. Using a resonance frequency meter of Hartmann and Braun, Einstein and de Haas could only measure frequencies at steps of a half cycle per second, as the device only was equipped with standard coils at certain fixed frequencies. To interpolate to the intermediate frequencies, they used an ammeter to measure the current provided in the generator. The ammeter therefore became their only measure of frequency between the frequencies given directly by the frequency meter.

As the frequency was changed, the double excursion lengths were measured by eye as the light beam oscillated back and forth across a scale some 145 cm. from the mirror. By plotting their results, Einstein and de Haas found the resonance curve shown in Figure 6. From this graph, the following data was taken and reduced to a value for P from which the gyromagnetic ratio could be calculated. (See Figure 7.) After eliminating measurements in which they felt the excursions were too small to measure accurately,[28] the authors found that their experimental result ($2m/e = 1.11$ within a 10% error) was in excellent agreement with their theoretically predicted value ($2m/e = 1.13$).

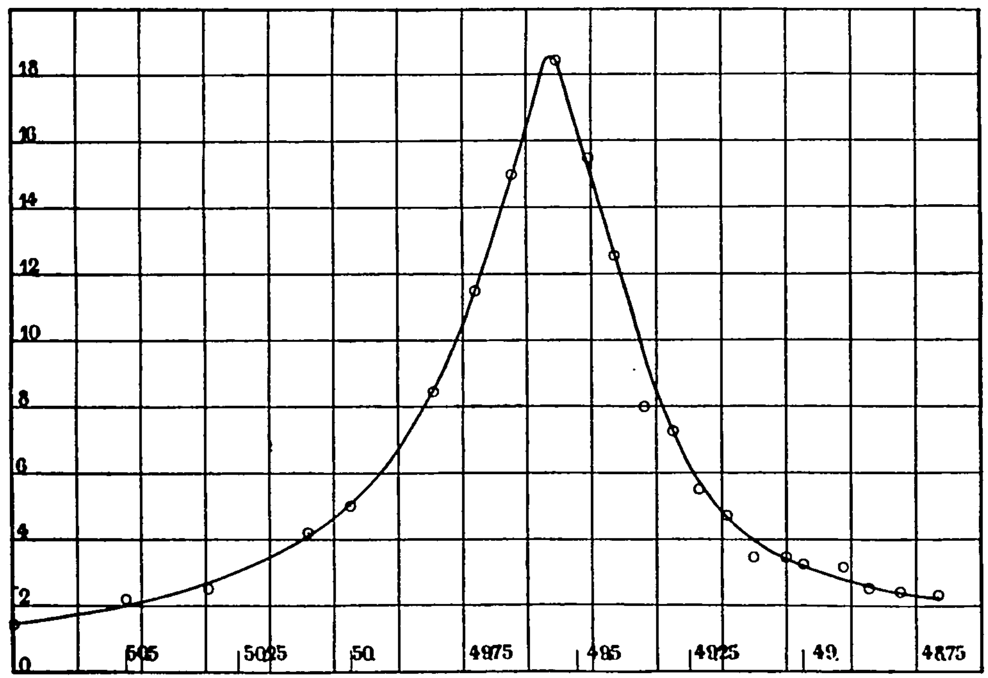
This is perhaps a good moment to establish a convention I will use below. Since $L/M = 2m/e$ was the original prediction for the gyromagnetic ratio for an orbiting electron, it has become standard to define the so-called g-factor by means of the following relation:

$$L/M = (2m/3)(1/g).$$ [29]

For an orbiting negative electron, g is therefore simply 1. One might expect g to be 1 for any orbiting system. However, this is not always true. For instance, a spinning classical sphere with mass distributed evenly throughout the sphere and charge distributed only on the surface would have a g-factor of 5/3. Indeed, by a suitable disposition of charge and mass, one could create a spinning classical sphere with any g-factor desired.

Einstein's theoretical prediction therefore corresponded to a g-factor of 1; his empirical result was equivalent to a g-factor of 1.02 with an error of .10. Such an extraordinary agreement with theory left the two physicists persuaded that they had verified Ampère's hypothesis. They concluded,

Figure 6. Data points from Einstein and de Haas' original paper. From Einstein and de Haas. "Experimental Proof" (ref. 27), p. 708. Vertical axis shows projection of light beam on scale in millimeters; horizontal axis shows frequency of oscillating magnetic field in cycles per second.



18
16
14
12
10
8
6
4
2
0
505
5025
50.
4975
495
4925
49.
4875

Figure 7. Numerical data used by Einstein and de Haas to find gyromagnetic ratio. From Einstein and de Haas, "Experimental Proof" (ref. 27), p. 710. The column marked "ordinates" shows the light beam deflection in millimeters; ν is the width of the resonance curve at height the beam deflection is measured; b is the fraction of maximum light beam excursion; $(b^2/(1 - b^2))^{1/2}$ is proportional to the damping constant of the fiber. For reference, the quantity in the right-hand column is inversely proportional to the g-factor.

Ordi-nates	ν	b	$\sqrt{\frac{b^2}{1-b^2}}$	$\nu\sqrt{\frac{b^2}{1-b^2}}$
15	0,0911	0,812	1,32	0,120
12	0,152	0,649	0,853	0,130
9	0,221	0,488	0,560	0,124
7	0,293	0,380	0,413	0,121
5	0,403	0,271	0,280	0,114
4	0,489	0,217	0,222	0,108
3	0,618	0,163	0,165	0,0957

> The precision of the agreement may be accidental as our determination must be taken to have an uncertainty of about 10 percent; however, it has been demonstrated that the results of the theory of the orbiting electron sketched at the beginning have been quantitatively established (at least approximately) by the experiment.[30]

It is interesting to note that if, contrary to what Einstein and de Haas did, one includes the other three data points (below 7mm) and calculates the g-factor, one obtains a number about five percent higher than their published result.

However, other sources of error enter at this stage as well, since as shown above it is necessary to determine I, α, and M. The maximum excursion was obtained by observing the light deflection at resonance; the moment of inertia was found by adding a cross bar of known moment of inertia and measuring the free-motion frequency with and without the bar. Unfortunately, the saturation magnetization was calculated in a way that may have introduced two important systematic errors. First, the hysterisis curves of the material were used to determine magnetization as a function of the solenoid's B-field. The iron rod may or may not have been similar in composition to the standard sample. Second, the solenoid's magnetic field itself was not measured, it was calculated from the constants of the coil. It is interesting to note in this regard that in their original article, Einstein and de Haas reported the volume integral of the cylinder as $M = 1260$ and the saturation intensity of magnetization as 458; in the English and Dutch papers that appeared shortly afterward, they more modestly eliminated the third significant figure and presented these two quantities respectively as M = 1300 and I_s = 470, writing their final result with one less significant figure (λ = 1.1 instead of 1.11).[31]

In addition to these sources of errors there are several other types of systematic errors present that Einstein and de Haas recognized: 1. if the axis of rotation does not correspond to the axis of the magnetic field, the suspended bar will acquire an alternating horizontal magnetization moment. This alternating horizontal magnetization can then couple with the Earth's magnetic field to produce a large disturbing effect at just the frequency of the Einstein-de Haas effect. 2. conversely, the horizontal component of the Earth's magnetic field can magnetize the iron bar directly. Then if there is an alternating horizontal magnetic field in the solenoid, another very strong disturbing effect will be introduced, also at the frequency of the Einstein-de Haas effect. Both of these disturbing effects could be several orders of magnitude stronger than the Einstein-de Haas effect. They therefore might easily mask the sought-for signal if the Earth's field is not adequately neutralized. Maxwell, much to his disappointment, had discovered this some fifty years before.[32]

Neutralization of the Earth's magnetic field became the most crucial and delicate aspect of the early failure and eventual success of this experiment. In their original experiment, Einstein and de Haas used hoops with a radius of one meter with coils wound around them to eliminate the Earth's magnetic field. The field strengths of the hoops were monitored by an ammeter measuring the current flowing through them.

To examine the field in the immediate vicinity of the rod, Einstein and de Haas used a galvanometer and a device which measured the induction of the Earth's magnetic field. As a final check on the compensation of the Earth field, they rotated the glass fiber and then turned on the current oscillators. When no further variation between the amplitude of oscillation from angular position to angular position was detected, the Earth field was considered eliminated.

In later experiments the above method proved too crude, and after de Haas returned to Holland on 1 April 1915,[33] both he and Einstein separately began to work on the problem of further eliminating the residual horizontal Earth field. De Haas set out to eliminate the first disturbing effect by directly wrapping the wire of the solenoid on the suspended rod. This assured the coincidence of the rotational and magnetic axes. Coupling still, however, could occur between the magnetized rod and the transverse Earth field. The Earth field thus still needed to be eliminated. To this end, de Haas arranged an array of permanent magnets. First, a large magnet was employed to eliminate the Earth field near the center of the bar, and then two smaller ones were used to compensate for the field near the poles. Any remaining field was compensated by a second coil placed at right angles to the rod-coil assembly. The two coils were attached in series and a variable resistor was placed in parallel to the horizontal coil. De Haas could then adjust the distance of the coil from the rod-coil assembly and regulate the resistor to neutralize the Earth field.[34]

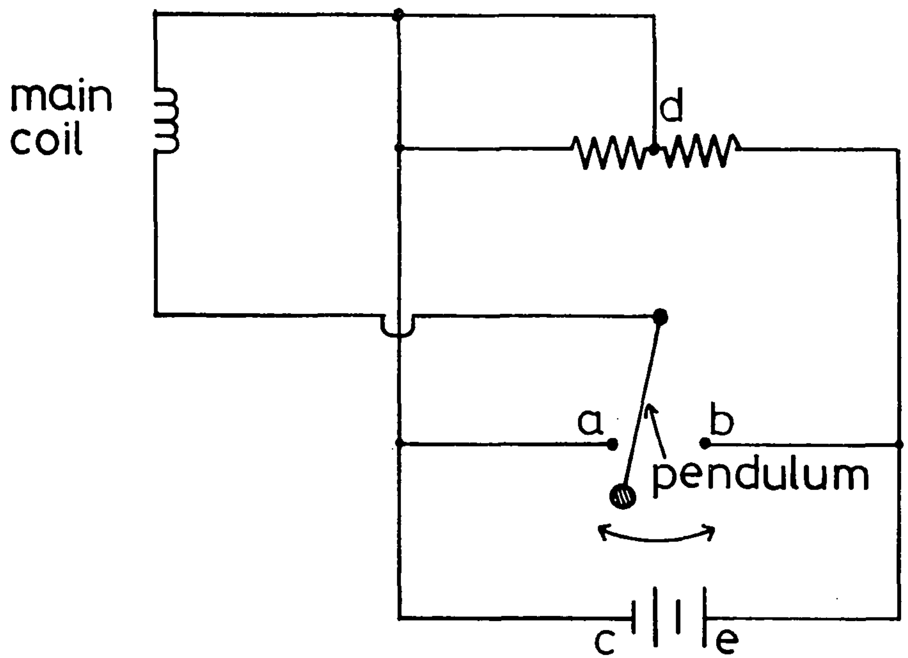
A final innovation of de Haas was to use a current pulse instead of a sinusoidal current by adapting a pendulum to complete a circuit each half cycle. (See Figure 8.) When the pendulum hits at a, a current pulse goes one way through the coil; when it hits b, the current pulse flows in the opposite direction. After a variety of control experiments, de Haas was able to determine that the deviations due only to the Einstein-de Haas effect corresponded to $g = 1.2$, from which he concluded: "This time again I have not had in view an accurate quantitative determination; yet it may be mentioned that the quantitative agreement between experiment and theory is quite satisfactory. At the same time a way is opened for a later accurate determination of e/m."[35]

By explicitly writing that he considered the method a valid way of deriving e/m, de Haas in effect had assumed the orbiting electron version of Ampère's hypothesis to be correct. Whereas in modern physics the coefficient of 2m/e is taken to be representative of nuclear spin, electron spin, or the orbital g-factor, for de Haas and many other workers at the time it was a foregone conclusion that ferromagnetism and paramagnetism were due to orbiting electrons. For the moment, however, de Haas put stress on the method and presented his quantitative results modestly.

In private de Haas had already begun to suspect that the difference between $g = .86$ and $g = 1.0$ was significant. Though de Haas's original letter does not survive, a recently discovered set of Einstein letters to de Haas reveals that de Haas was quite concerned about the discrepancy. Einstein replied,

> I am very happy to hear about your work on the effect. [Einstein is referring to the work of

Figure 8. Alternating current pulse circuit. De Haas designed this circuit to minimize the time the solenoid field had to be on, as it tended to disturb the magnetized cylinder after the reversal of magnetization had taken place. When the pendulum hits a the current flows along cade. When the pendulum hits b the current flows along ebdc. A pulse therefore is produced of the form shown above.


main
coil
d
a
b
pendulum
c
e

> de Haas just described.] I also have conducted experiments, in which I eliminated the remanent magnetization through the discharge of a capacitor. The experiment won't work yet because despite the short duration of the field (10^{-3} seconds), strong vibrations of the little cylinder set in, hiding the effect. This is naturally avoided with your method. I can believe that your 10% discrepancy with the theory is real. If this is so, however, then it would be very significant.[36]

The experiments of his own that Einstein mentioned here were written up a short time later and received for publication in February 1916 as a "Lecture Experiment."[37] Einstein's idea, as mentioned in the letter cited above, was to demagnetize rather than reverse the magnetization of the iron rod. For this, a much smaller B-field was required. Like de Haas, Einstein used an alternating pulse rather than a sinusoidally varying current.

After adjusting the glass fiber such that the rod naturally oscillated a a second or two per cycle, the experimenter noted the deviation of the light marker. Each time the beam reached a maximum, he pressed a key pulsing the circuit. This would either markedly amplify or brake the swing, thereby demonstrating at least qualitatively the effect looked for. Again, Einstein made reference to the problem of compensating for the Earth's field and aligning the rod properly, but no specific details were given nor any quantitative results published.

Thus, after constructing at least four different versions of the experiment, Einstein and de Haas were convinced that they had verified Ampère's hypothesis with orbiting electrons serving as the "current whirls." Qualitatively, all four experiments pointed to a gyromagnetic effect; after some confusion (corrected by Lorentz), the phase relations between current and oscillation clearly indicated the charge on the electrons to be negative. Finally, after two separate quantitative determinations, results were found that can be expressed as:

$$g = 1.02 \pm .10 \quad \text{(Einstein-de Haas 1915)}$$

$$g = 1.16 \quad \text{(de Haas 1916).}$$

It is crucial to remember, however, that de Haas, even in his later work, took his measurements in principle to be a measure of 2m/e, and not a multiplicative constant by which this quantity was to be multiplied.

Einstein's and de Haas's work was soon being discussed in a great many places; from Princeton and Ohio to Zurich and Uppsala, a variety of physicists began to focus their attention on the confirmation or refutation of Einstein's experimental claims.

3. Einstein's Theoretical Preoccupations

Einstein must have had strong reasons for taking up an experimental problem in 1915. He also seems to have been persuaded that he would find that g was equal to unity. What was Einstein's theoretical motivation?

Very early in their joint paper, Einstein and de Haas inserted a very compressed paragraph addressing this question. They wrote that if electrons orbit around atoms, then by Maxwell's equations they should radiate their energy away. This, they asserted, "is surely not the case," and continued,

> Furthermore it follows from the Curie-Langevin Law that the magnetic moment of the molecule is temperature-independent. Therefore, since the magnetic moment still exists at $T = 0$, there should remain an energy associated with the motion of the orbiting electrons at $T = 0$. Many physicists understandably resist the acceptance of this so-called "zero-point energy."[38]

Einstein's abbreviated remark goes to the heart of his motivation to perform the experiment. Pierre Curie had discovered experimentally in 1895 that in paramagnetic substances (substances whose magnetization are proportional to the applied field), the magnetic susceptibility varies with the reciprocal of the temperature.[39] In 1905, using the statistical techniques of Boltzmann, Curie's colleague Paul Langevin derived the Curie Law by assuming that each atom carried a magnetic moment, m, owing to the circulation of electrons.[40] From these assumptions Langevin found the suspectibility to be

$$\chi = m^2 N/(3kT),$$

where N is the molar density, k Boltzmann's constant, and T the temperature. For Einstein, Langevin's success in predicting the Curie law gave credence to the assumption that there existed a temperature-independent atomic magnetic moment. Einstein hypothesized that tnis atomic magnetic moment might be due to Ampèrean current loops composed of circulating electrons. Since such electronic motion would persist at a temperature when all molecular motion ceased, it fell in the class of so-called "zero-point energies." But electronic motion was just one of several possible mechanisms by which energy could remain in a collection of atoms at absolute zero. For instance a "zero-point energy" might be due to molecular vibrations.

Einstein's concern with the zero-point energy began before his interest in this aspect of Langevin's work, dating back to his work in 1907 on the quantized harmonic oscillators of specific heat theory. Later, in 1911, the zero-point energy was taken up again in yet another context when Planck used it in his "second theory."[41] In the new theory, oscillators were allowed continuous absorption of energy but discontinuous emission. Only when an oscillator had acquired an energy equal to a multiple of $h\nu$ can it emit a light quantum. Using these assumptions, Planck claimed the average energy of an oscillator included an additional term equal to $h\nu/2$, present even at absolute zero. Planck, however, paid no further attention to this energy as he thought the frequency should be independent of temperature, and therefore the term would not contribute to the specific heat.[42] Since the zero-point energy seemed in 1911 to be a necessary consequence of the latest version of Planck's quantum theory, it was natural for Einstein to enquire into the possibilities of deriving experimental consequences of this additional energy.

In 1913, in collaboration with Otto Stern, Einstein began such an enquiry.[43] In the first part they pointed out that the rotational motion of a molecule should, by statistical mechanics, depend on temperature. They therefore created a model for diatomic hydrogen for which they could compare predictions for specific heat with and without the assumption of the zero-point energy. Their collaborative work was, in a sense, an indirect continuation of Einstein's 1907 analysis of the specific heat associated with a system of quantized harmonic oscillators. After this paper appeared, Nernst had proposed that Einstein's quantization be extended to rotational as well as vibrational motion. Thus Einstein was continuing an old interest when he and Stern quantized the rotational energy of a molecule by setting the average rotational energy of a molecule,

$$E = J/2(2\pi\nu)^2,$$

equal to the Planck expression for the average energy of an oscillator of frequency ν. Their expression,

$$E = J/2(2\pi\nu)^2 = h\nu/(\exp\,(-h\nu/kT) - 1),$$

referred to a collection of molecules all rotating at the same frequency in equilibrium with the radiation.

In order to determine whether or not the zero-point energy should be included, the authors calculated the specific heat from the above equation by eliminating ν and forming $c = dE/dT$. They then compared the resulting equation with and without an additional term $h\nu/2$ on the right hand side. The two equations yielded different expressions for specific heat as a function of temperature that could be compared to the experimental data of A. Eucken. Having graphed the two theoretical predictions against the experimental data, Einstein and Stern concluded, "Eucken's results on the specific heat of hydrogen make probable the existence of a zero-point energy of $h\nu/2$."[44] Thus far Einstein and Stern's argument was based on the Planck radiation law and therefore on the quantum hypothesis. In the second part of their article they reversed their approach. By assuming a zero-point energy Einstein and Stern contended that no further demands of discontinuity were needed to derive the Planck radiation law. Einstein doubted, however, that "other difficulties" (that he did not specify) could be conquered without the assumption of quanta.[45] The Einstein and Stern paper thus provided a double argument for the existence of one kind of zero-point energy. However, the inadequacy of this first part of Einstein and Stern's work was soon revealed in a critique by Ehrenfest,[46] who made the more realistic hypothesis that the molecules had a statistical distribution of rotational frequencies. By doing so, he showed rotational energy did not lead to a specific heat formula in good accord with experiment and concluded that the Einstein-Stern attempt to justify a zero-point energy was not valid.

When Einstein and de Haas began their experiments, they therefore knew that at least part of Einstein's earlier argument for the existence of a zero-point energy had collapsed and that a new one was needed. In a letter to his friend Michele Besso, Einstein reported on his soon to be completed experiments in glowing terms:

> ...the experiment will soon be finished. It will also have proved the existence of a zero-point energy. A wonderful experiment, too bad you can't see it. And how devious [heimtückisch] Nature is, if one wants to approach it experimentally! I've gotten a longing for experiment in my old age.[47]

Another consideration relating to the quantum may also have played a role in motivating Einstein to conduct the experiment. In 1913 Niels Bohr published his first paper on quantum theory in which he accounted for the Pickering lines in terms of orbiting electrons.[48] Soon after the paper appeared, Einstein hailed Bohr's work as "one of the greatest discoveries."[49] Since orbiting electrons were precisely the object of the gyromagnetic experiment, Einstein may have hoped to provide an indirect confirmation of Bohr's theory.

Einstein's interest in the various phenomena and theoretical developments associated with the quantum problems and the zero-point energies may, in part, have motivated his interest in the Einstein-de Haas experiments. But in keeping with a scientific style that had characterized much of his writing and thinking for over a decade, Einstein did not start the Einstein-de Haas paper by listing specific experimental anomalies or theoretical inadequacies. Instead, as he often did, he began the Einstein-de Haas paper by underlining a theoretical asymmetry, in particular a lack of unity among certain physical explanations. Einstein's interest in testing Ampère's hypothesis began shortly after his work on special relativity and before much of his and others' work on zero-point energies. It may be that considerations of unity played an important role in his choice of an experimental problem as it had in his selection of theoretical ones.

Gerald Holton[50] has stressed the importance such considerations played in Einstein's thought both in his formulation of the Special and General Theories of Relativity. Einstein's 1905 special relativity paper begins not with a problematic experimental result but with the sentence, "It is known that Maxwell's electrodynamics--as usually understood at the present time--when applied to moving bodies leads to asymmetries which do not appear to be inherent in the phenomena." The asymmetry in Maxwell's equations (as they were understood before Einstein) was evident, for example, in the explanation of what occurred as a conducting coil and magnet approached each other with a velocity $\vec{v}$. In the rest frame of the conducting coil, the magnet is moving at a velocity $\vec{v}$, and so the magnetic field is changing. According to Maxwell's equations, a changing magnetic field is associated with the production of an electric field. In this case, the electric field would produce a current in the conducting coil. In the rest frame of the magnet, however, the conducting coil is moving with velocity $\vec{v}$. Again, according to Maxwell's electrodynamics, a static magnetic field produces a current in a moving conductor. Hence from the two frames there are different explanations that predict the same current to be found in the conducting coil. Thus two explanations were offered for what seemed to Einstein to be a single phenomenon since the magnitude and direction of the current produced is predicted to be the same from either frame. Einstein's relativity theory is essentially the introduction of a unified theoretical representation of this single phenomenon.

Einstein's theory of General Relativity was another of his attempts to unify disparate explanations. In Newtonian physics inertial mass and gravitation both entered the theory as fundamental concepts. By contrast one of the founding principles of General Relativity was that there be only one type of mass in physics. Einstein's search for symmetry and unity affected his thought about radiation and statistical mechanics as well. Martin J. Klein has pointed out that one of Einstein's objections to the wave theory of radiation, as it had stood since the time of Maxwell, was an asymmetry in the process of the emission and absorption of light quanta.[51] In 1909, Einstein argued[52] that on one hand, a single electron suitably displaced could generate an expanding spherical electromagnetic wave, thus producing the emission of radiation. On the other hand, a great many emitters would be needed to create a collapsing spherical wave to produce the absorption of radiation by a single electron. It was partially in an effort to restore the symmetry of absorption and emission that Einstein, in 1905, had introduced the light quantum into physics. Again, one explanation replaced the two different accounts of absorption and emission.

Einstein's search for a unifying principle was another motivation for him to verify Ampère's hypothesis, as is evident from the first section of Einstein and de Haas's 1915 paper:

> Since Oersted discovered that magnetic effects are produced not only by permanent magnets but also by electrical currents, there have been two seemingly independent mechanisms for the generation of a magnetic field. This state of affairs itself brought the need to fuse together two essentially different field producing causes into a single one--to search for a single cause for the production of the magnetic field. In this way, shortly after Oersted's discovery, Ampère was led to his famous hypothesis of molecular currents which established magnetic phenomena as arising from charged molecular currents.[53]

For Einstein, the state of affairs of having two essentially different causes for what seemed to be one phenomenon was already a powerful argument for the search for a single cause.

There is, however, more to Einstein's interest in and commitment to the experimental verification of the gyromagnetic ratio as $2m/e$. For in the second paragraph, the authors claim:

> Also the electron theory (especially as it has been developed by H. A. Lorentz), is tied essentially to Ampère's hypothesis in the demand for a unified conception of the production of electromagnetic

> fields. According to the electron theory, however, the molecular currents, as in general all electrical currents, are made from moving elementary charges.[54]

Again Einstein referred to a unification, this time to the Lorentz electromagnetic theory. Einstein later amplified on Lorentz's contribution, contending that before Lorentz's electromagnetic theory was developed, physicists treated the electric and magnetic fields as conditions governing matter.[55] Thus the electric field and the dielectric displacement were treated as independent entities. By contrast in Lorentz's scheme these fundamental vectors of the electric and magnetic fields act on the electrons which by their rearrangement affect the total field by the contribution of their own fields.[56] For Einstein at least, his experiment was a test of two fundamental unifying principles: the Ampèrean hypothesis and the Lorentz electron, as well as being an examination of various quantum hypotheses.

4. Barnett: From Terrestrial Magnetism to Einstein's Error

There was another path to the gyromagnetic experiments, almost entirely separate from the one that began with Ampère's hypothesis. This other path commenced with terrestrial magnetism, one of the oldest mysteries of physics. Though the attempt to link the Earth's magnetic field to gyromagnetic effects would be all but completely abandoned a few years later, it served as the motivating factor for many theoretical and experimental investigations.

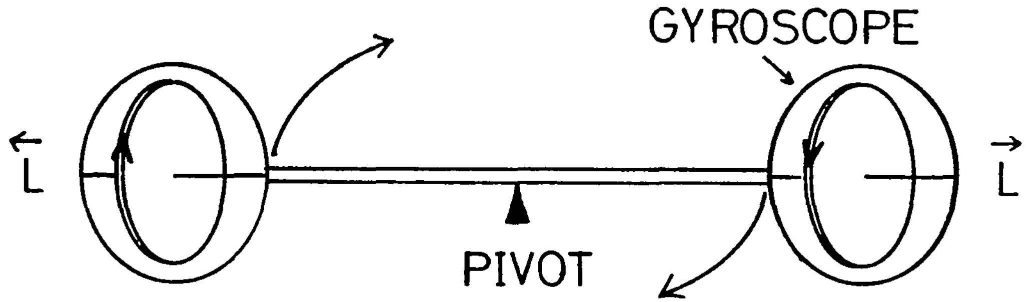
In an 1890 lecture on spinning tops given to the British Association, John Perry speculated on the connection between rotation and magnetization.[57] Perry likened the spinning molecules he took to compose matter to a "honeycombed mass with a gyrostat in each cell." This, he asserted, was no chance analogy. Magnetized matter might be nothing else but the state of iron (for instance) in which all the microscopic gyrostats were oriented. This suggested an experiment. If one gave an unmagnetized piece of iron an angular acceleration the little spinning molecules that composed it should experience a torque tending to orient them. Rotation should therefore produce magnetism. Though unsuccessful in his attempts to induce magnetism in this way, Perry attributed his "failure to the comparatively slow speed of rotation which [he] ...employed, and to the want of delicacy of... [his] magnetometer."[58]

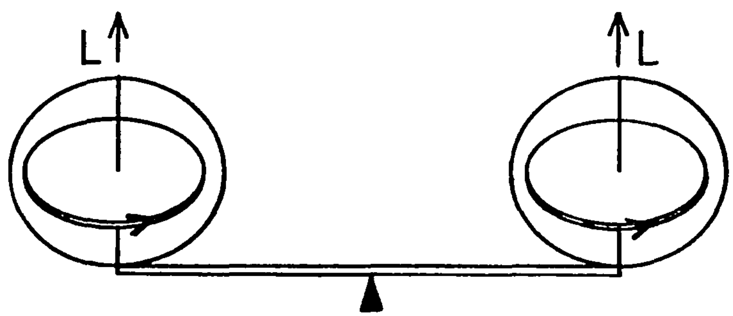
In 1909 Samuel J. Barnett at Ohio State University proposed a similar connection between rotation and magnetization while thinking about the relation of the Earth's magnetic field and its rotation.[59] Barnett hypothesized that magnets were composed of many oriented atomic or molecular systems with individual magnetic moments. If the atomic systems had negative electrons orbiting around positive centers, one could predict the direction of a magnetic field that would result if the iron were given an angular acceleration. The following mechanical analogy may help explain why the electronic current loops would be expected to orient themselves. Suppose a bar is placed on a fulcrum, with a gyroscope placed on either end (see Figure 9). The two gyroscopes have angular momenta equal in strength but opposite in direction. As in the Einstein-de Haas effect, this system of zero total angular momentum is analogous to the randomly oriented current loops inside a bar of unmagnetized iron. If the bar is given an angular acceleration about the fulcrum, the gyroscopes will orient themselves to conserve total angular momentum.

Barnett performed his first measurements on a steel rod that was accelerated quickly from zero to ninety rotations per second. Using a ballistic galvanometer, he could measure the magnetic field produced by the spinning rod. For these first attempts, he reported a field of 1/1500 gauss, with a sign associated with the presence of orbiting negative electrons.

However, unlike Richardson (and later Einstein and de Haas), Barnett was <u>not</u> especially interested in the consequences of his experiments for Ampère's hypothesis, Lorentz's electron theory, or zero-point energies. Instead, Barnett's conclusion addressed his main concern, terrestrial magnetism. "This effect, if substantiated by later work, will account for a minute part of the earth's magnetism, but

Figure 9. Mechanical analogy to Barnett effect.


GYROSCOPE
L
L
PIVOT


L
L

apparently, for only a minute part."[60] Barnett did not make quantitative theoretical calculation of the expected magnetic field from a collection of orbiting electrons, since his primary interest was in the Earth's rotationally generated field. Not until after reading Richardson's 1908 paper did Barnett associate his effect in theory with orbiting electrons. It is therefore anachronistic to assign a g-factor to his result; for later discussion, however, the magnetic field he recorded in 1909 is equivalent to g = 11.

During the years following 1908 in which Barnett continued to examine rotational magnetization, he was not alone in his curiosity about the connection between the Earth's rotation and its magnetic properties. The most important other work was the Presidential Address of Arthur Schuster to the Physical Society of London on 9 February 1912, entitled "A Critical Examination of the Possible Causes of Terrestrial Magnetism."[61] "We know," Schuster declared, "that the earth behaves like a magnet with its axis inclined at an angle of about 12 degrees to the geographical axis of the earth. Is this near coincidence between the two axes merely accidental?" Not surprisingly, the new president concluded that there was a physical connection between the two vectors.

In support of his own suggestion that rotation causes magnetization, Schuster examined various candidates for the source of terrestrial magnetism. First, he rejected the idea of a magnetized core of the Earth since iron should lose its magnetization "even with the most modest estimate of the internal temperature of the earth."[62] Still unknown effects of high pressure on the critical temperature of iron might save his hypothesis, and for this reason the possibility of a magnetized iron core was left open. Second, Schuster turned to and rejected the view that terrestrial magnetism might be caused by a massive rotating current. Such a current would rapidly be dissipated and there was no evidence either of a cause of the current nor of earlier magnetic fields much more powerful than those found today. Finally, Schuster dismissed the idea that an external magnetic field might induce a magnetic moment on the Earth since there is no evidence of such extraterrestrial magnetic fields. As final evidence for his hypothesis that rotation causes the magnetic moment, Schuster cited the secular variation of the magnetic north about the geographical pole. This, he added, could be explained if the electrons responsible for the Earth's magnetic field were free to precess about the geographical pole.

By 1914, Barnett had read Schuster's speech, and more importantly had realized the connection between his work and the attempts of Maxwell and Richardson to measure the gyromagnetic effects that would follow from Ampère's hypothesis. Adapting Maxwell's equation for the torque on a circular wire with a current through it, Barnett showed that he could expect theoretically a magnetic field to be produced at the pole of the iron cylinder equal to

$$H/n = -7.1 \times 10^{-7} \text{ gauss/change in rps [corresponds to } g = 1.0\text{]}.$$ [63]

In Barnett's 1915 experiments, he replaced the ballistic galvanometer with a fluxmeter. In addition, he improved the sensitivity of the measurement by adding a "compensating bar" identical to the rotating bar. A coil is wound around the compensating bar in the opposite direction from the coil surrounding the test bar. By keeping the compensating bar at rest, the circuit automatically compensates for any flux changes due to extraneous fields such as the motor's. (See Figure 10.)

After calibrating the fluxmeter and compensating for the Earth's magnetic field with several large coils, the experiments were undertaken. Barnett's result, as printed in the conclusion, was that H/n was less than half of the expected result for orbiting electrons. If we again anachronistically convert this into a g-factor:

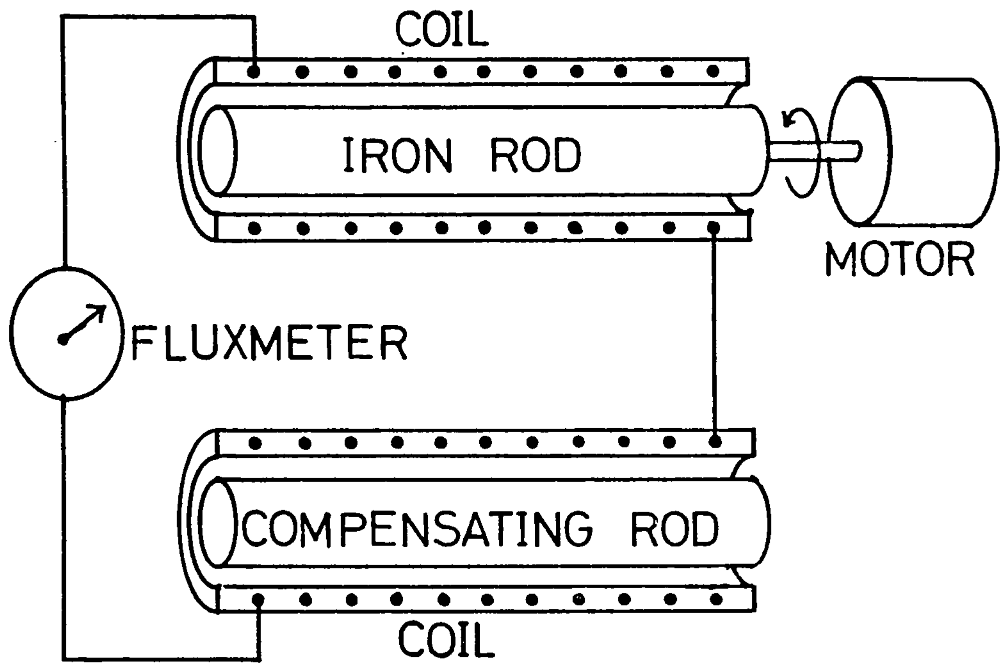
$$g = 2.3.$$

From the Earth's angular velocity and magnetization, it follows that the effect would amount to less than 10^{-10}th of the Earth's magnetic field. On the face of it such a discrepancy would seem to dash all hopes of using rotation to explain the terrestrial magnetism. But, implicitly drawing on Schuster's speech,[64] Barnett ended the body of the paper with the remark that conditions inside the Earth might explain the increased magnetization needed to account for terrestrial magnetism.[65]

Barnett mentioned two other effects (centrifugal displacement and thermionic displacement of electrons) that might also come into play in the creation of the Earth field. Yet these last vague hopes were never elaborated. By the time Barnett published again two years later, he had abandoned his interest in terrestrial magnetism and the subject was never again given prominence in Barnett's papers, even as the source of his original idea.

Barnett's move away from terrestrial magnetism undoubtedly came

Figure 10. Schematic illustration of Barnett's 1915 experiment.


COIL
IRON ROD
MOTOR
FLUXMETER
COMPENSATING ROD
COIL

when he read the Einstein and de Haas work of 1915. This was but shortly after his own paper of that year appeared in print. There can be no doubt that the paper had a great influence on the American experimentalist. For contrary to the conclusion of his (1915b) paper that was specifically designed to explain terrestrial magnetism, Barnett opened the 1917 paper with the following words:

> Before these [Barnett's 1914-1915] experiments were made only one method of magnetizing a body was known, viz., placing it in a magnetic field. These experiments not only revealed another and entirely new method, but they also confirmed the fundamental assumptions on which the results had been predicted: They proved...(1) that Ampèreian currents, or molecular currents of electricity in orbital revolution, exist in iron; (2) that all or most of the electricity in orbital revolution is negative; and (3) that it has mass, or inertia, so that each orbit behaves like a minute gyrostat...[66]

Instead of terrestrial magnetism, for the first time Barnett introduced "Ampèreian currents," following the lead of Einstein and de Haas.

The 1917 experiments were based on a new experimental apparatus where the fluxmeter has been replaced by a magnetometer. In the old method the motor was turned on and off; the fluxmeter measured the resulting change of flux. With a magnetometer (essentially a suspended coil with a current running through it) the magnetic field can be measured directly since the deflection of the coil or magnet will be proportional to the strength of the field. The magnetometer was more sensitive than the old ballistic galvanometer but it was also much more susceptible to outside disturbances.[67] Barnett therefore took special care to compensate for the Earth's magnetic field (p. 12), shifts in the rotor's altitude (p. 18), longitudinal motion of the rotor (p. 19), temperature variation (p. 20), and mechanical vibrations (p. 20). Despite these careful precautions, Barnett's results were not, as one might think, much closer to the now accepted value near g = 2. In 1917, Barnett took the proximity of his new result to g = 1 as a confirmation of his measurements' validity.

The first clue to Barnett's new outlook comes at the beginning of his 1917 paper, when Barnett recounts his early experiments. After introducing an equation describing the expected result if the current is due entirely to the orbit of a negative electron (g = 1), Barnett wrote: "If positive electricity also participates [g should be larger]. The mean value of...[g] obtained in my 1914 experiments was...[2.0]; and [g] was found to be independent of speed within the limits of the experimental error."[68] Now this is a rather extraordinary remark. Barnett's articles written in 1914 and published in 1915 (1915a,b) only reported a result equivalent to g = 2.3. No data equivalent to g = 2.0 were ever presented in reduced form. The only data Barnett could have been referring to are the raw data from his 1914 experiments displayed in (1915b) as "Table 1. Earlier Observations and Results," where he reported that the "weighted mean differential deflection per unit speed...equals 0. 057mm per revolution per second."[69] If we reduce this mean, it is equivalent to g = 2.0. The reason Barnett himself had not reduced these raw data is clear from the sentence following Table 1:

"After the completion of the work thus described it was decided to repeat the rotations in a region in which the Earth's [magnetic] intensity was still more completely annulled." To this remark he added the following footnote: "The desirability of this course was realized from the first and was also mentioned by Dr. Rosa at the Philadelphia meeting of the Physical Society, Dec., 1914."[70] Thus from Barnett's perspective in 1915 the g = 2.0 data were not reliable and the g = 2.3 data were trustworthy. Naturally, at the end of the body of the paper, in the conclusion, and in his 1915a paper Barnett only reported the g = 2.3 result. Why, then, did Barnett reduce apparently unreliable data for publication in the introduction to his 1917 paper?

There seem to be two reasons for this. First, Barnett was concerned about receiving credit for his discovery, as is evident in almost all of his publications over the following thirty years. By pointing to data gathered in 1914, Barnett made it clear that his results predated the Einstein-de Haas publication of 1915. But this would not explain why he left out the 1915, g = 2.3 result from the introduction of his 1917 paper. The mystery, however, becomes much clearer in the conclusion to the 1917 paper in which Barnett reported that his new magnetometer results ranged from g = 1.4 to g = 1.1. He concluded: "The differences are in the same direction as in the earlier experiments on iron, which gave...[g = 2.3 and 2.0 instead of g = 1]." Here for the first time it is evident that Barnett now expected the result g = 1, as he continued,

> ...but the experimental errors on account of the great difficulties involved, are such that importance cannot in my opinion be attached to the discrepancies. The investigation must be taken as confirming equation (1) [g = 1.0] both qualitatively and quantitatively on the assumption that only electrons are in orbital revolution in the molecules of all the substances investigated.[71]

This position represents a change of Barnett's point of view. Whereas before Einstein and de Haas he had concluded that a g-factor above 1.0 indicated the presence of orbiting positive charge, now "no importance can be attached to the discrepancies." The sudden change during these two years is clearly due to the impression the Einstein-de Haas papers have made on him. This explains: 1. why he dropped all reference to terrestrial magnetism; 2. why he began his paper with a discussion of Ampèreian currents; 3. why he in 1917 expected to get g = 1.0; 4. why he dropped the g = 2.3 and stressed the g = 2.0 result (since it is closer to g = 1.0); and most importantly, 5. why in 1917 he concluded that he had "qualitatively and quantitatively" confirmed the equation g = 1.0. The question that jumps to mind, however, is how was his <u>data</u>, g = 1.4 to g = 1.1, influenced by his new theoretical predispositions? Let us defer this question for a moment.

5. The Revenge of Experiment: Stewart, Beck and Arvidsson

After Richardson's unsuccessful attempt to measure the gyromagnetic ratio in 1908, there were several attempts at the Princeton laboratory to refine the experiment. Finally, in 1915, John Quincy Stewart and Maurice Pate began a series of investigations,[72] that may have been encouraged by Barnett's completion of the converse experiment. One problem that had plagued the Princeton group was one that had led Einstein and de Haas separately to restructure their original technique: as soon as the suspended rod became magnetized, it interacted directly with the solenoid in such a way as to mask completely the searched-for effect.

De Haas had dealt with the problem by wrapping the solenoid directly around the suspended iron and using short pulses to reverse the magnetization. Einstein had attacked the difficulty by using weak, extremely short pulses delivered at the resonant frequency to demagnetize the bar.

Stewart pursued Einstein's idea. In addition to the remanent magnetism method, Stewart introduced three fundamental improvements. First, he designed a system of six square coils arranged on the faces of a cube centered on the suspended sample. Each pair of facing coils was wired in series. In this way, the Earth's magnetic field would be effectively eliminated: first in a rough way by the ratio of coil turns between the vertical and the horizontal coils, and then in a fine way by adjusting the currents through the coils.

Second, Stewart made use of narrower and longer wire samples than had Einstein and de Haas. This minimized the amount of demagnetization that took place due to the action of the poles on the rest of the sample. Finally, Stewart cleverly employed two exploring coils to eliminate the transverse magnetization of the rod (this was due to a small permanent magnetization and a longer transverse magnetization induced by the Earth's and the solenoid's fields).

By a suitable arrangement of the exploring coils, Stewart could find the magnetic moment of the sample and eliminate it by demagnetizing the sample in small increments until the moment began to be reversed. Then he could measure the free-swinging period of the sample, assured that no magnetic control was being exerted on it by the Earth's field. When the bar was re-magnetized, a different period was observed; the compensating coils then were adjusted so the free period was achieved again. When the free period was restored, Stewart considered the Earth's transverse field neutralized. Similarly, the solenoid's transverse field was eliminated by magnetizing the sample and adjusting the solenoid's inclination until the free-swinging period was restored.

Once the disturbing effects were eliminated, Stewart experimentally determined the smallest demagnetizing current which was still effective, and performed the measurement. (See Figure 11.) His result, averaged over sets of experiments on nine different wires (but excluding sets where the wires were above a certain thickness), was:

$$g = 2.0 \pm 0.2.$$

One test of the accuracy of Stewart's experiment was that when demagnetization took place from a downwards magnetization, an opposite

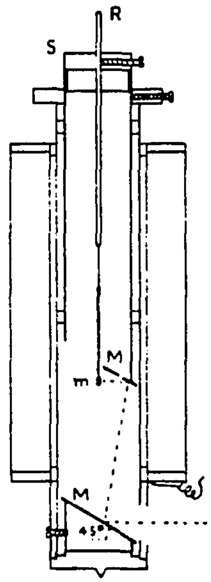


Figure 11. Stewart's 1918 Apparatus. (From Physical Review, 11 (1913), p. 102.

but equal displacement of the light beam was observed from that observed when the original magnetization was upwards.

Stewart accepted the result that g is approximately 2, and concluded that one of two possibilities must hold. Either 1. only negative electrons are rotating, but they do not fully react on the bulk matter (slippage hypothesis), or 2. positive and negative charges are in rotation in opposite directions. Stewart's readiness to turn to a theory that included orbiting positive as well as negative charges may stem from his work with Richardson, as he reproduced Richardson's general formula (mentioned in Section 1) for the ratio of angular momentum to magnetic momentum for both positive and negative charges.

Ultimately, Stewart dismissed the first possibility (slippage) as being unlikely in light of the coincidence between his results and those of Barnett. Consequently, he concluded positive charge must be rotating as well. He added:

> According to Sir Ernest Rutherford's theory of atomic structure, all the positive charges are concentrated in a very small "nucleus" at the center of the atom, while about half the negative electrons are rotating around the nucleus at distances very large compared with its diameter.[73]

Taking a proton to electron mass ratio of 1850, and using the measured gyromagnetic ratio, Stewart concluded that "the angular velocity of the rotating positive nucleus is about equal (but opposite in sign) to that of the inner ring of electrons."[74]

Meanwhile in Zurich, another experimentalist, Emil Beck, set out to repeat Einstein's experiment with more precision.[75] Unlike Stewart, Beck continued to use Einstein and de Haas' resonance method, reversing the magnetization of the iron cylinder with an oscillating magnetic field. Beck was sufficiently confident both in Einstein's orbiting electron theory and in his own measurements to have written, "In the opinion of the writer this method lends itself very well to an exact determination of the important quantity, e/m." Only a strong belief that $L/M = 2m/e$ allowed him to say he was measuring e/m and not simply the gyromagnetic ratio, L/M.[76]

Three improvements over Einstein and de Haas's experimental method gave Beck this confidence. The first and most important was the elimination of the awkward frequency measuring system employed by Einstein and de Haas. Instead of making a few measurements with a resonance meter and interpolating between the points by varying the current to the motor, Beck developed his own frequency measuring device that gave very accurate measurements of small frequency differences. To do this, he exploited one of the disturbing effects in Einstein's experiment: direct coupling takes place between the magnetized rod and the horizontal component of the alternating magnetic field, causing torsional oscillations. In Beck's device, a coil is wound parallel to a suspended permanent magnet in series with the main solenoid. This causes a strong oscillating horizontal field which forces the suspended magnet to oscillate on its fiber. For any given frequency, by adjusting the length of the wire attaching the magnet to the support structure, Beck could find the length corresponding to resonance (maximum excursion length). In this way, very small frequency differences could be measured with great accuracy. In turn, an accurate measure of frequency differences led to a more accurate determination of the damping constant, and therefore of the gyromagnetic ratio.

Beck's second innovation was to use a photographic plate to record the deflections of the light reflected from the little mirror mounted on the rod. This gave him an additional, and direct measurement of the damping constant P, which was determined by setting the rod in free oscillation and reading from the developed film the decaying amplitude of the excursions.

Thirdly, Beck had a much better determination of the constants B (magnetic field), Q (the moment of inertia of the rod), and I_s (the magnetization intensity of the rod), which enter in the calculation of g. To obtain the magnetic field inside the solenoid, he used a tiny mirror-galvanometer suspended by a wire. (By contrast, Einstein and de Haas had only calculated this quantity.) He also obtained a much better correspondence between his calculated and measured moment of inertia for the rod. Finally, Beck <u>measured</u> the saturation magnetization of the rod by wrapping a coil around the rod and attaching it to a calibrated galvanometer. When the magnetic field was suddenly turned on, the rod became magnetized causing a change in the magnetic field. Since the ambient magnetic field from the coils had already been determined, the meter deflection could be used to determine the magnetization of the rod.[77] When both resonance and photographic measures for the gyromagnetic ratio were calculated and averaged, Beck obtained a result corresponding to

$$g = 1.8,$$

a number outside the limits of Einstein and de Haas's error bar. The discrepancy led him to check "all causes of error" and to review with special care the alterations he had made from Einstein and de Haas' original procedure.[78]

Before Beck published his work, Einstein came to visit him in Switzerland where they discussed the experiment. Impressed by Beck's work, Einstein reported to de Haas, "In Zurich a really good experimentalist (Herr Beck) has repeated our measurements on the torque exerted on a ferromagnet and only found a half of the theoretically expected effect...."[79] Beck hesitated to announce that he had made a new determination of m/e. Instead, he concluded that either 1. there was a new type of a electron, 2. the nucleus or positive particles were circulating in the opposite direction from the electrons, or 3. the situation was somehow more complicated than previously suspected.

Beck's and Stewart's results were soon confirmed by G. Arvidsson, working independently at Uppsala.[80] Like Beck, Arvidsson referred to his measurements as a determination of m/e, and he too used the method of resonance by reversing magnetization of the iron cylinder. But as Arvidsson had not yet seen the results of Beck or Stewart, to him the discrepancy between his result and Einstein's was somewhat worrisome. After presenting his data, which averaged to

$$g = 2.12,$$

Arvidsson concluded, "In my opinion, one must acquire a more exact knowledge of phenomena involving statistical magnetization in an oscillating field before we can say anything precise about the results."[81]

With the measurements of Stewart, Beck and Arvidsson all pointing towards g = 2, the simple model of orbiting electrons was cast into doubt. By the time of the April 1921 Solvay conference, the issue was of considerable concern to many of those interested in the physics of the electron. De Haas reported on his experiments at the meeting, and there followed a discussion that included Lorentz, Richardson, and Larmor.[82]

Like Barnett, de Haas succumbed to the temptation of resurrecting earlier, unreliable data and presenting them later along with his final results. Speaking of his experiments with Einstein, de Haas wrote, "The numbers we found for 2m/e in our experiments were...[g = 1.4] and ...[g = 1.0]. The second value was almost the classical value... [g = 1.00] which led us to believe that experimental errors had made the first too...[large]."[83] This first result of g = 1.4 which de Haas reported came from a set of experiments explicitly rejected by Einstein and de Haas in their 1915 paper. After presenting the calculated and observed double deflections of the light marker, they had not calculated the experimental 2m/e, and for good reason. For immediately after the deflections were presented, they added the caveat that to satisfy the conditions specified in the theoretical calculation, it was necessary to have an almost instantaneous reversal of magnetization. For their first experiment this was not the case, Indeed it was principally this factor which led Einstein and de Haas to repeat the experiment. It was therefore not an altogether accurate representation of the earlier experiments for de Haas to present the two pieces of data g = 1.0 and g = 1.4 on equal footing, though this probably reflects his growing conviction that Stewart, Beck and Arvidsson might be right.

After the conference, de Haas published two new sets of data, whose averages were:

March 1921: g = 1.55

July 1921: g = 1.11.

Then, after discussing various disturbing effects, de Haas concluded:

> The other authors cited in this report found double the classical value of e/m. As for me, I am tempted to consider the exact value of the effect per se as still an open question. Be that as it may all the observers found a value of e/m which was too large. A part of this torque is therefore disappearing and escaping our observations. The idea was presented that a positive nucleus turning at a high speed could absorb a part of the torque. But this hypothesis seems to me far-fetched and unlikely; I think instead that if the bases of the theory are unimpeachable, other hidden movements must be considered.[85]

Shortly later, in a discussion of the problem at the Reichsanstalt in Berlin, Einstein repeated de Haas's dissatisfaction with the increasingly well-accepted value of g as 2. He asked, "Can't we investigate exactly the magnetic rotation effect here in the Reichsanstalt?

There still remains no certainty over the numerical factor [g]."[86]

In October of 1920, after the results of Stewart, Beck and Arvidsson had been published, Barnett submitted another article. This time he presented a brief report indicating that his 1917 work might be defective. Eddy currents, Barnett remarked, had been detected in copper samples when used in place of iron in the magnetometer experiments. "This probably accounts for at least a part of the discrepancy between the results obtained by the two methods [1915 galvanometer and 1917 magnetometer experiments-P.G.]."[87] He could now assert:

> All the rods gave values about [twice] [g = 1] instead of [g = 1], or even less, as in the experiments on iron... indicating an effect of positive electricity or else indicating that negative electricity alone is involved, but has for the motions responsible for magnetism, a smaller value of m/e than that determined in known experiments.[88]

Barnett repeated these beliefs later that year.[89]

By 1922, Barnett had prepared an article on his new research for the Bulletin of the National Research Council.[90] There, he stressed his 1915 results (g = 2.3 and 2.0). And in another bit of revisionist history, his results of 1917 disappeared with the words, "In 1917 we completed an investigation of steel, cobalt, and nickel by a magnetometer method, and obtained values of [g] which were, as before, all negative, and whose means were intermediate between the values previously obtained for steel and twice those values." (This translates to: g was between 1 and 2.) No numbers were given.

To explain the new (or rather old) result that g = 2, Barnett left the Einstein orbiting electron theory to invoke the theories of W. Voigt and M. Abraham. As Abraham had shown, if one takes the charge of an electron to be spread evenly over the surface of a sphere and calculates the mass purely electrodynamically, the ratio $L/M = m/e$ for rotational motions corresponds to g = 2. For a rotating electron with the charge distributed through its volume, a result is obtained equivalent to g = 5/14.[91] From these suggestive numbers, Barnett concluded that either 1. positive electrons or "magnetons" are in rotation, or 2. that one of the two rotating electrons suggested by Abraham is responsible for the effect, or 3. that a new kind of "magneton," different from the orbiting electron is responsible for these gyromagnetic effects. Although Barnett had no results of his own to report, he was thoroughly convinced that his original 1915 results were correct (in agreement with Stewart, Arvidsson, and Beck) and that the 1917 results were spurious.

In 1922, Barnett felt under siege. Louis Bauer, the head of the Carnegie Institution Department of Terrestrial Magnetism, became involved in a protracted feud with Barnett. Among other issues, Barnett's single-minded commitment to small improvements in his old experiment infuriated Bauer.[92] At the same time, Barnett's laboratory assistant was complaining that Barnett would not let him undertake any but the most mechanical and routine tasks or "have any part in observations or reductions concerned with the experiments under way."[93] Even the instrument makers began to despair over the possibility of improving the apparatus in the way Barnett sought. By 1922, over one-seventh of the instrument makers' time was devoted to Barnett's experiment alone. Finally, J. A. Fleming (the Assistant Director of the laboratory) wrote to Bauer recommending that no further work should be committed to Barnett's experiment by the instrument makers. Fleming concluded his letter by reporting:

> In my judgment the mechanical difficulties which Dr. Barnett is trying to overcome in the existing apparatus arise from fundamental mechanical defects...if...they could be temporarily improved...the adjustment probably would not be permanent and might not hold long enough even for any extended, reliable series of observations.[94]

Partly as a result of these pressures, Barnett finally left the laboratory and continued his work at Cal Tech with his old equipment. It was thus in California in 1925 that Barnett and his wife, L. J. H. Barnett, finished a massive 89-page study of the Barnett effect with a detailed study of gyromagnetic effects with an exhaustive discussion of errors. As a example, a few of the sections' headings are:

39. Eddy current effects of the lower magnetometer magnet
40. Effect of air currents on bedplate
43. Elimination of thermal effects on magnetometer
47. Error due to thermal effects of journal friction on magnetization
51. Errors from axial displacement of the rotor
53. Error from the Thomson repulsion effect
54. Errors from mechanical disturbances
55. Errors due to inequality of right-handed and

left-handed speeds

It is clear that the Barnetts did not want to be fooled again. After 159 sets of observations, they presented their result: g = 1.89, and concluded with the assertion, "We do not see how the error can be greater than 2%."[95]

With this much more precise data in hand, Barnett abandoned his ideas on the Abraham electron and turned to a quite different field of physics then under intense debate.

> Our phenomenon is undoubtedly connected closely with the Zeeman effect, as our magnetons may be considered to be executing regular precession upon them brought about by the rotation....As Landé has suggested, the anomaly in the Zeeman effect, which Sommerfeld and Debye had partially explained by the ideas of spatial quantization (now supported in the field of magnetism by the work of Pauli, Sommerfeld, Epstein, Gerlach, and Gerlach and Stern), is probably related to the anomaly in our phenomenon. This anomaly Landé and Sommerfeld have attempted to explain by a process which appears to be equivalent to identifying our magnetons with the atoms in the s-state and attributing to this a value of [g] equal to m/e [g ≐ 2] which is approximately the value of [g] given by our experiments.[96]

Thus, once again Barnett changed the theoretical analysis of his experiment. This time he identified his experiment with the spectroscopic phenomena which would shortly be explained by Goudsmit and Uhlenbeck as deriving from electron spin. However, the actual disovery of spin was made quite independently of the gyromagnetic experiments, according to Uhlenbeck.[97]

If theoretical interpretations still remained vague in 1925, at least the quantitative determination of g was becoming increasingly accurate. Shortly before Barnett's 1925 work, two English physicists, A. P. Chattock and L. F. Bates, used a modification of Stewart's experiment to obtain a value of g as:

$$g = 1.97.$$ [98]

Their apparatus was then further refined by C. N. Sucksmith and L. F. Bates to obtain

$$g = 1.99 \pm .024.$$ [99]

Barnett too investigated the Einstein-de Haas effect, and in 1931 obtained a value of:

$$g = 1.929 \pm .006.$$ [100]

Many other variations on the gyromagnetic experiments have since been performed, especially on paramagnetic substances, but perhaps the most exact have been those of G. G. Scott, working at the Research Laboratories of the General Motors Corporation. One of the more recent publications on the subject appeared in 1962 when Scott reported his best determination of g to be:

$$g = 1.919 \pm .002.$$ [101]

By this time the spin-orbit and orbit terms were known to be very dependent on the properties of specific substances and therefore the g-factor in itself revealed little of fundamental importance to physics. The gyromagnetic experiments had long since passed from the forefront of physics.

6. Summary and Conclusions: On Theoretical Predispositions

As can be seen in Figures 12 and 13, this episode in the history of experimental physics is a rather extraordinary tale. First, Maxwell and Richardson both failed to get any result at all in attempting the gyromagnetic experiments. Then, unaware of their research, Barnett began work on the converse effect: when rotating an iron rod, he detected a magnetic field which was more than five times the strength current physics predicts he "should" have found. After reading Richardson's paper and revising his own experiment, Barnett in 1925 arrived at a value approaching g = 2.3 and was quite satisfied with the explanation that positive ions were orbiting in the atom opposite to the negative electrons. His main conclusion, however, was that this effect, combined with unknown conditions at the center of the Earth, might make it possible that the Earth's rotation was the cause of terrestrial magnetism.

Almost simultaneously, Einstein who (unlike Barnett) had very strong reasons to believe that g = 1, performed the experiments in 1915 with de Haas. For here was a chance for Einstein to confirm Lorentz's electrodynamic theory, Langevin's explanation of the Curie law, Planck's zero-point energy hypothesis, and Ampère's molecular current hypothesis. After at least four different experimental apparatus had been constructed, Einstein and de Haas seemed to have conclusively verified the theory that orbiting electrons were responsible for permanent magnetism. They determined that g = 1.02 ± .10, and in a second quantitative series of experiments the following year, de Haas found

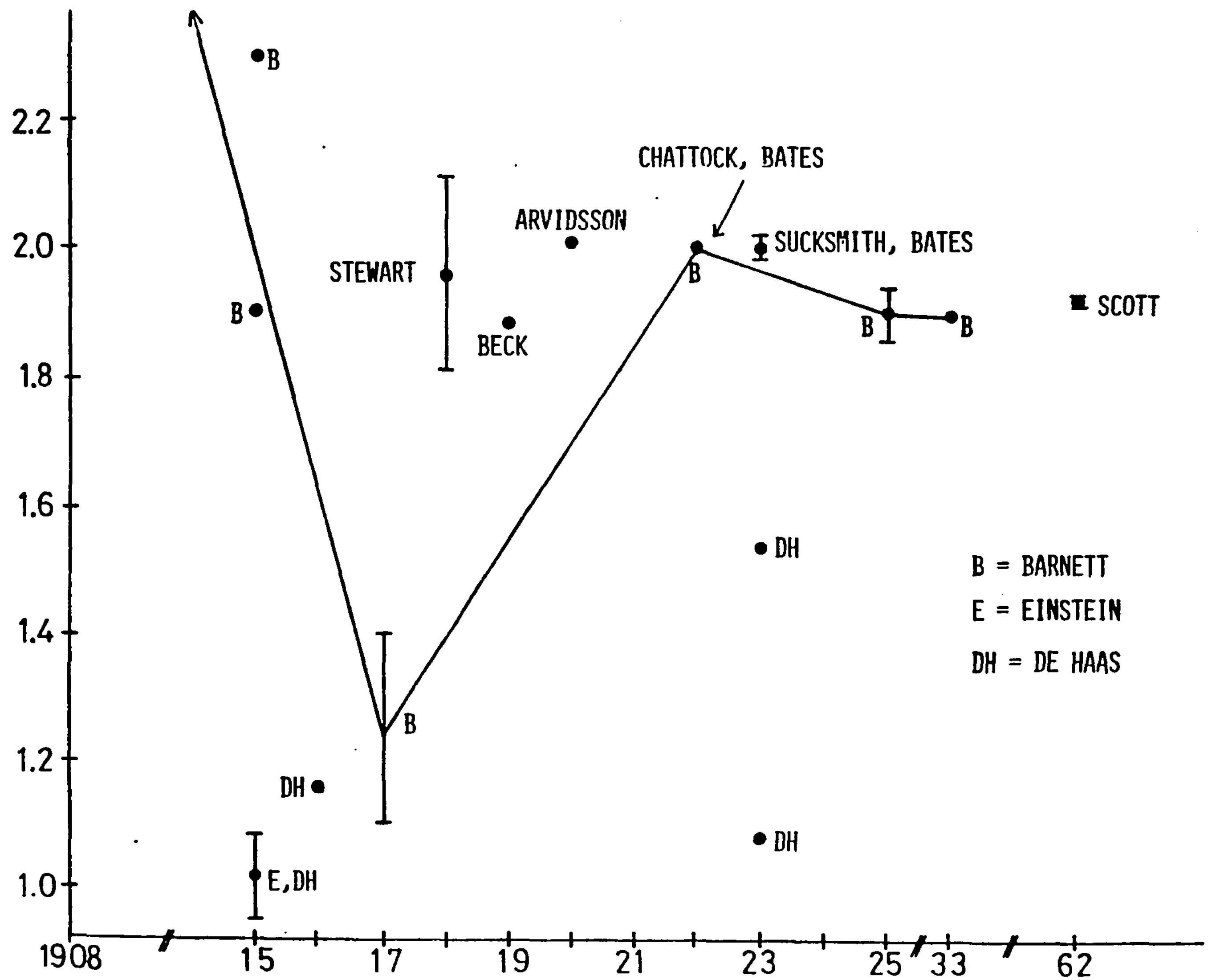
Figure 12. Summary of Gyromagnetic Results.

Experimenter	Place	Publication Date	Results (g-factor)
Barnett	Physical Lab., Ohio State University	1915	1.9 (1914) 2.3 (1915)
Einstein-de Haas	Physikalisch-Technische Hochschule (Berlin)	1915	1.2 ± .10 (1.45--dismissed)
de Haas	Teyler Institute	1916	1.16
Barnett	Ohio State	1917	1.39 to 1.09 (within error of 1.0)
Stewart	Palmer Lab., Princeton	1918	1.96 ± .15
Beck	Eidgenossische-Technische Hochschule (Zurich)	1919	1.88
Arvidsson	Physical Institute (Uppsala)	1920	2.12
Barnett	Carnegie Institute (Washington)	1922	"approximately 2"
de Haas	Teyler Institute	1923	1.54 (March 1921) 1.08 (July 1921)
Chattock & Bates	University of Bristol	1922	1.97
Sucksmith & Bates	University of Bristol	1923	1.99 ± .024

Figure 12, continued.

Experimenter	Place	Publication Date	Results (g-factor)
Barnett	Cal Tech	1925	1.89 ± .04
Barnett	Univ. of Calif. at L.A.	1931	1.929 ± .006
Scott	General Motors Laboratories (Michigan)	1962	1.919 ± .002

Figure 13. Publication year vs. g-factor. The solid line traces Barnett's results as a function fo time.



PUBLICATION YEAR VS. g-FACTOR

g = 1.16.

Then Barnett, obviously influenced by Einstein's theory and experiment, repeated his own work and concluded that he too had vindicated the orbiting electron theory: g was somewhere between 1.4 and 1.1. However, the story was far from over, as three experimentalists, working independently, soon determined that g was not equal to one. Stewart, Beck, and Arvidsson each published a quantitative result nearer to twice the Einstein value. Within months Barnett published again, asserting that he too believed that g was approximately 2. In the two years that followed, he improved his result, abandoned Einstein's theory, and adopted one of Abraham's electron theories to explain his result of g = 1.89.

Meanwhile, de Haas (1921) repeated his work in two additional series of experiments, now aware that at least four other researchers were finding a g-value near to twice his original one. At the Solvay meeting, de Haas reported a g-value of 1.54 where he asserted that he still considered the value of g to be an open question. After the meeting he repeated his experiments for the last time: his cumulative result of g = 1.08 was only a few percent different from his original result with Einstein six years before. The next year, in Berlin, Einstein too maintained that the value of g was still open to question. During this time Barnett refined his method further, and in 1925 published a massive paper with an average result of g = 1.929 ± .006. By 1933 the Dirac theory was well known; Barnett was then able to attribute his result to a complex interaction of spin and orbit effects.

Among the explanations for the way theoretical predispositions influence experimental results is the one given by Thomas Kuhn in the article "The Function of Measurement in Modern Physical Science."[102] His argument is essentially as follows: the measurements necessary to test new theories often bear on phenomena at the limit of our experimental capabilities. As a result, relative to the size of the effects searched for, random error is very great. This leaves open the possibility for experimentalists and theorists to interpret the necessarily ambiguous results as confirming their theory. Had more precise techniques of measurement been possible, these same results could just as easily have confirmed an opposing theory. For example, Kuhn cites the case of Laplace's prediction of the speed of sound in air. In this episode, Laplace arrived at a theoretical prediction in excellent agreement (a discrepancy of only 2.5%) with the experimental results of Delaroche and Berard.[103] Their result, though, now seems to differ by over 40% from modern measurements and theory. Kuhn concludes that any measurement like that of Delaroche and Berard must also fit other theories, "and it is only within the experimental spread covered by the phrase 'very nearly' that nature proved able to respond to the theoretical predisposition of the measurer."[104]

In other words, the collection of relevant data has a sufficient spread or "scatter" (as Kuhn calls it elsewhere[105]) that competing theoretical explanations may both be compatible with the experimental results.[106] However, Kuhn's explanation only applies if <u>in retrospect</u> we can see that all and only "relevant" data have been used. Sometimes irrelevant data can be excluded at the time of the experiment. Thus, for instance, in Stewart's experiments, he excluded the thick wires from his average value for g because of demagnetizing effects he felt were systematically distorting the results. Similarly, Barnett excluded his 1914 experiments because he realized upon their completion that the Earth's field had not been adequately neutralized. Unfortunately, at the time of the experiment it is not always possible to identify which data are "relevant" and which must be discarded.

In light of this I would like to suggest a somewhat different interpretation of the way theory influences the outcome of experiment, one that depends neither on Gestalt-like mis-seeing nor on the large spread of random errors. First, it is crucial that Einstein and de Haas had a theoretical belief--that the current whirls were orbiting electrons--which translated into a definite quantitative prediction. In addition, the measurements under investigation were extremely delicate: the movement of oscillating reflected light beam from the Einstein-de Haas cylinder is on the order of millimeters and the Barnett effect depends on the production of a magnetic field of the order of 10^{-5} gauss. But most crucially, as a result of systematic errors from a variety of sources, the mean result was shifted in different directions often without leaving tell-tale traces of large dispersion in the results as would random errors. De Haas, in a publication of 1923, reported:

> As to the largely dis-crepant values found by us and by myself I must remark that these experiments were made in a very short time, and that we were glad already to detect the effect in an unobjectionable way. The numbers serving

> for the calculation of the effect were but roughly known. So we did not measure the field of the magnetizing coil, we calculated it; moreover the coils were wound rather irregularly and not made for the purpose of the experiment. Also we did not measure the magnetization of the rod, we estimated it. We mentioned all this in our original paper. These preliminary results seemed to us rather satisfactory, and it will be easily understood that we were inclined to consider the value $g = 1.02$ as the better. (1921 Solvay Conference)

None of these errors would cause a spread in the results (in fact, Einstein and de Haas gave their probable error as 10%, far from including $g = 2$). Furthermore, such an explanation nowhere requires us to take recourse to the world of Gestalt images.

Barnett, too, later tried to explain his 1917 $g = 1$ results. In 1922 he wrote:

> A long suspected systematic error has been found in the 1917 magnetometer observations, causing the results to differ considerably from those obtained by the method of electromagnetic induction in 1914 and 1915 is now fully confirmed.[107]

As in the case of the original Einstein and de Haas measurements, the result now considered correct lies outside the range of Barnett's 1917 data ($g = 1.1$ to 1.4). Some of the other systematic errors later pointed out by Barnett included such seemingly harmless elements as trolleys passing outside, incomplete compensation for the Earth's magnetic field, and expansion of the rod during rotation.

The quantitative expectation which Barnett had in 1917 was undoubtedly reinforced (as it was for de Haas, Beck, and Arvidsson) by the frequent interchanging in his writing of the two sides of the equation L/M and 2m/e. The measured quantity L/M thus became inseparable from the prediction that this should be 2m/e. In the minds of these experimentalists, they were measuring the gyromagnetic ratio associated with an Ampèrian current whirl, rather than testing Ampère's hypothesis.

It is not enough, however, to say that theoretical predispositions are a purely pernicious factor. In the case of Maxwell, for instance, it is precisely his <u>lack</u> of a quantitative prediction (because he had no orbiting electron model) that left him with no idea how big an effect he was looking for in his "second experiment." Had he known, as de Haas and de Haas-Lorentz showed much later, that he could expect a tilt of the apparatus of only .00013 radians, he would never have used this experiment as evidence for the non-inertial nature of current. Similarly, it may well have made it <u>more</u> difficult for Barnett in 1908 to have found the effect he was looking for because he had no quantitative prediction of the order of magnitude of the strength of the field he could expect.

The experimentalist would therefore seem to be in a continual dilemma. On the one hand, without a theory, one has no guiding quantitative prediction; the experimentalist is thus unlikely to find the effect at all, or to be able to dissociate it from disturbing effects. On the other hand, given a quantitative prediction, the experimentalist is eventually forced to declare (at least implicitly) that here are no more systematic errors. This "stopping place" is, naturally enough, often the predicted result. An experimental nuclear physicist, Martin Deutsch, once put the conundrum as follows:

> It is of course the ambition of every experimenter performing this kind of experiment to make a discovery, to sail safely between the Scylla of intellectual prejudice which makes us reject evidence not readily integrated without preconceived notions, and the Charybdis of irrelevance which has swallowed many working days spent in pursuit of instrumental artifice.[108]

In the series of experiments discussed here, the Scylla was the orbiting electron theory, and the Charybdis included the transverse magnetization of the rod by the Earth's field, the eddy currents of Barnett, and the improperly centered magnetic rod.

In light of what has been said here, one might expect that in experiments where both strong theoretical predispositions and a definite quantitative prediction are present, it will often be the case that the experimenter will find the result looked for whether or not this corresponds with what is later found to be the case. One might look, for instance, at some of the other famous experimental factors of two that have arisen in modern physics: parity violation, and the bending of starlight by the sun, or any number of cases where a new theory disagreed only slightly from the old in its quantitative prediction. In at least some of these I would expect compensating systematic errors to place the prediction and measurement in harmony and to place now current experimental results outside of earlier experimental error.

APPENDIX ON MANUSCRIPT SOURCES

Reference Number	Date	From Einstein to:
Postcards		
P-1	(1909)	Lorentz
P-2	17.3.15	de Haas
P-3	28.4.15	Lorentz
P-4	6.7.15	de Haas
P-5	9.7.15	de Haas and G.L. dHL
P-6	2.18.15	de Haas and G.L. dHL
P-7	7.8.15	de Haas and G.L. dHL
P-8	10.8.15	de Haas
P-9	14.8.15	de Haas and G.L. dHL
P-10	8.10.15	de Haas and G.L. dHL
P-11	12.5.24	de Haas and G.L. dHL
Letters		
L-1	3.2.15(?)	Lorentz
L-2	undated	de Haas
L-3	(April-May 1915-dHL)	G. L. de Haas-Lorentz
L-4	24.7.15	H. u. F. de Haas and G.L. dHL
L-5	(Summer 1915-dHL)	H. u. F. de Haas and G.L. dHL
L-6	(August 1915-AJK)	H. u. F. de Haas and G.L. dHL
L-7	(Fall 1915-dHL)	H. u. F. de Haas and G.L. dHL
L-8	(undated)	de Haas
L-9	9.5.19	de Haas

Reference Number	Date	From Einstein to
L-10	19.1.20	Lorentz
L-11	12.12.23	de Haas and G.L. dHL
L-12	17.10.24	de Haas and G.L. dHL
L-13	12.6.27	Lorentz
L-14	31.3.27	H. u. F. de Haas (with printed 50th anniversary reply)
L-15	11.4.29	de Haas
L-16	19.9.32	de Haas
L-17	2.5.33	de Haas and G.L. dHL
L-18	8.2.34	de Haas
L-19	21.12.53	de Haas
L-20	7.1.54	de Haas

NOTE: AJK = date provided by A. J. Kox; dHL = date provided by G. L. de Haas-Lorentz. G.L. dHL = G. L. de Haas-Lorentz.

In addition, we have found two letters from Ehrenfest to H. A. Lorentz, four letters from Ehrenfest to de Haas, eight letters from H. Haga to H. A. Lorentz, and 150 letters from H. A. Lorentz and/or his wife, mostly dealing with non-scientific matters.

A copy of the Einstein letters has been sent to the Einstein Archives in Princeton. The originals will remain in Holland where they will also be placed in an archival collection.

The Barnett letters all come from the personnel files under his name at the Carnegie Institution, Washington, D.C.

NOTES

[1]André Marie Ampère, "Mémoire Présenté à l'Académie royale des Sciences, le 2 Octobre 1820, où se trouve compris le résumé de ce qui avait été lu à la même Académie les 18 et 25 septembre 1820, sur les effets des courans électriques," Annales de Chimie et de Physique, 15 (1820), pp. 59-76 and pp. 170-218, on pp. 74-75. All translations are the author's. A partial translation is given in R. A. R. Tricker, Early Electrodynamics: The First Law of Circulation (Oxford: Pergamon Press, 1965).

[2]Ampère, 1820 (ref. 1), p. 76.

[3]Ampère, "Mémoire sur la théorie mathématique des phénomènes électrodynamique uniquement déduite de l'expérience, dans lequel se trouvent réunis les Mémoires que M. Ampère a communiqués à l'Académie royale des Sciences, dans les séances des 4 et 26 décembre 1820, 10 juin 1822, 22 décembre 1823, 12 septembre et 21 novembre 1825," Mémoires de l'Académie Royale des Sciences de l'Institut de France, VI (1823), issued 1827, pp. 175-388, on p. 303.

[4]J. C. Maxwell, Treatise on Electricity and Magnetism (Oxford: Oxford University Press, 1881).

[5]Ibid., pp. 202-203.

[6]Robert John Strutt, Life of John William Strutt (Madison: Univ. of Wisc. Press, 1968), p. 46. Letter from Maxwell to Strutt dated 18 May 1870.

[7]Maxwell, Treatise (ref. 4), p. 203.

[8] *Ibid.*, p. 205.

[9] W. J. de Haas and G. L. de Haas-Lorentz, "Een Proef van Maxwell en de Moleculaire Stroomen van Ampère," *Amsterdam Koninklijke Akademie Verslag Wissen Naturkuunde*, *24*, *1* (1915), pp. 398-404, on p. 404.

[10] Maxwell's work on absolute units was part of a project he undertook with Sir William Thomson, J. P. Joule and F. Jenkin to produce a series of reports on electrical standards. See Appendix D to report of 26 August 1863 in *Reports of the Committee on Electrical Standards appointed by the British Association for the Advancement of Science*, edited by F. Jenkins (London: E. & F. N. Spon, 1873).

[11] *Ibid.*, p. 111.

[12] Maxwell, *Treatise* (ref. 4), p. 206.

[13] *Ibid.*

[14] For more on the relation of current to momentum and energy after Maxwell, see Jed Z. Buchwald, "Matter, The Medium and Electrical Current: A History of Electricity and Magnetism from 1842 to 1895," Unpublished PhD dissertation, Harvard University 1974. Also I would like to thank J. Buchwald for helpful comments.

[15] The history of the Lorentz electron and the experiments associated with it have been discussed in many places. For instance see: R. McCormmach, "H. A. Lorentz and the Electromagnetic View of Nature," *Isis*, *61* (1970), pp. 459-497; A. I. Miller, "On Lorentz's Methodology," *British Journal for the Philosophy of Science*, *25* (1974), pp. 33ff.; K. Schaffner, "The Lorentz Theory of Relativity," *American*

Journal of Physics, 37 (1969), pp. 498-513.

[16]O. W. Richardson, "A Mechanical Effect Accompanying Magnetization," Physical Review, 26 (1908), p. 248.

[17]Ibid., p. 252.

[18]Ibid.

[19]O. W. Richardson, The Electron Theory of Matter (Cambridge: Cambridge University Press, 1914), p. 397.

[20]Ibid., p. 252.

[21]Max Flückiger, Albert Einstein in Bern (Bern: Verlag Paul Haupt, 1974), p. 172.

[22]Ibid.

[23]Prof. Dr. H.-J. Treder, "A. Einstein: 'Einfache Methode zum Nachweis der Ampèrschen Molekularströme," Wissenschaft und Fortschritt, 2 (1979), p. 53.

[24]Dieter Hoffmann, "Albert Einstein und die Physikalisch-Technische Reichsanstalt," Wirkung von Albert Einstein und Max von Laue. Akademie der Wissenschaften der DDR, Institut für Theorie, Geschichte und Organisation der Wissenschaftlichen Kolloquien. Heft 21. (Berlin, 1980), pp. 90-102 on p. 90.

[25]A. Einstein, "H. A. Lorentz, His Creative Genius and His Personality," in H. A. Lorentz Impressions of His Life and Work, ed., G. L. de Haas-Lorentz (Amsterdam: North-Holland, 1957), pp. 5-9, on p. 8. Originally appeared as "H. A. Lorentz als Schöpfer und als

Persönlichkeit," Report No. 91 from Rijksmuseum voor de Geschiednis de Natuurwetenschappen, Leiden, June 1953. See also Martin J. Klein, Paul Ehrenfest, v. 1 The Making of a Theoretical Physicist (Amsterdam: North-Holland, 1970), p. 300.

26 Hoffman, "Einstein" (ref. 24), p. 91.

27 Einstein and de Haas, "Experimenteller Nachweis der Ampèreschen Molekularströme," Verhandlungen der Deutschen Physikalischen Gesellschaft, Berichte 13, 17 (1915), pp. 152-170 on p. 152. This was the original publication of their joint work. All translations are the author's. The paper was then changed and republished in English and Dutch: (a) "Experimental Proof of the Existence of Ampère's Molecular Currents," Royal Academy of Amsterdam, Proceedings, 18 (1916), pp. 696-711; (b) "Proefondervindelijk bewijs voor het bestaan der moleculaire stroomen van Ampères," Koninklijke Akademie van Wetenschappen Te Amsterdam Urslagen der Afdeeling Natuurk, 23 (1914-1915), pp. 1449-1464. I have found the article by V. Ia. Frenkel, "Kistorii effekta Einshteina-De Gaaza," Uspekhi fizicheskikh nauk, 128 (July 1979), pp. 545-557 to be very helpful. In it Dr. Frenkel stresses (as I do here) that Einstein was involved in the experiments and was not a passive onlooker. See also: Sir Edmund Whittaker, A History of the Theories of Aether and Electricity (New York: Humanities Press, 1973), pp. 243-5.

28 Einstein and de Haas, "Experimental Proof," (ref. 27), p. 710.

29 Because a confusing variety of units and conventions are used in the gyromagnetic experiments to be discussed here, I have

systematically translated or converted the results to the experimental value for the g-factor they implied so they may be compared. Where the form of the original result has been important, this will be made clear in the text. Furthermore, all results that are converted to g-factors will be presented with the same number of significant figures and their corresponding error bars, if they are in the original.

[30] Einstein and de Haas, "Experimenteller Nachweis" (ref. 27), p. 170.

[31] *Ibid.*, p. 169; Einstein and de Haas, "Experimental Proof" (ref. 27), p. 711.

[32] The following numbers illustrate how overwhelming the sources of systematic error can be. The torque due to the Einstein-de Haas effect, T_{EdH}, can be estimated from the saturation magnetic moment of the iron cylinder, $M = 470$ ergs/gauss and the frequency of oscillation of the magnetic field, $\omega = 50\ \text{sec}^{-1}$. The change in angular momentum during one reversal of magnetization is ΔL:

$$\Delta L \sim (2)(5\cdot 10^{2}\ \text{ergs/gauss})(1\cdot 10^{-7}\ \text{gm-cm/esu-sec}) = (10^{-4}\ \text{gm-cm}^{2}/\text{sec})$$

so $T_{EdH} = \omega\Delta L \sim 5\cdot 10^{-3}$ ergs.

The disturbing effects, by contrast, exert torques that are much larger.

1. If the oscillating iron cylinder is 1% off alignment with the solenoid, there will be a magnetization in the horizontal direction of approximately 10 ergs/gauss coupling to a transverse Earth field

which, if uncompensated, could be on the order of 1 gauss. This would represent a torque due to misalignment, T_M:

$$T_M = (10 \text{ ergs/gauss})(1 \text{ gauss}) = 10 \text{ ergs}$$

2. Conversely, if the Earth transverse field is on the order of 1 gauss, it could magnetize the iron bar, which has a magnetic susceptibility of approximately 2×10^3 $1/\text{cm}^3$. A horizontal magnetization of the iron cylinder will result on the order of 2×10^3 ergs/gauss, which will couple to the horizontal component of the solenoid's alternating field. Supposing the horizontal component of the solenoid's alternating field to have been 1% of its total, or 0.5 gauss, we obtain a torque due to interaction with the Earth's field, T_E:

$$T_E = (2 \times 10^3 \text{ ergs/gauss})(.5 \text{ gauss}) = 10^3 \text{ ergs}$$

Both disturbing effects are therefore several orders of magnitude stronger than the Einstein-de Haas effect.

In their original paper two other sources of error are discussed and defined by Einstein and de Haas. These are:

1. Eddy currents which were known not to exist by repeating the experiment using a conducting, but non-magnetizable material of the same dimensions as the iron cylinder.

2. Permanently magnetized crystals within the bar which by their components in the horizontal direction might not be reversed by the ambient field. They could then couple to the horizontal Earth field or the oscillating horizontal solenoid field.

[33]Treder, "Einstein," (ref. 23).

[34]Yet a third disturbing effect was discussed by de Haas: If there is some hysteresis it is possible for the horizontal magnetization to be non-parallel to the horizontal field during part of the current cycle. If the cycle itself is non-symmetric it may come to pass that the lag and lead domains do not compensate and therefore provide a net torsional disturbance. De Haas, "Further Experiments on the Moment of Momentum Existing in a Magnet," Royal Academy of Amsterdam, Proceedings, 18 (1916), pp. 1281-1299.

[35]*Ibid.*, p. 1282.

[36]Unpublished letter from Einstein to de Haas labeled "Autumn 1915" by G. L. de Haas-Lorentz (de Haas' wife) at a later time. This letter is part of a collection of letters I located in Holland with the assistance of A. J. Kox. They had been among the papers left by G. L. de Haas-Lorentz. I would like to thank Hendrik Antoon Lorentz for making the letters available. A copy of the collection has been deposited with the Einstein archives in Princeton.

[37]A. Einstein, "Ein einfaches Experiment zum Nachweis der Ampèreschen Molekülarströme; von A. Einstein," Deutsche Physikalische Gesellscahft, Verhandlungen, 18 (1916), pp. 173-177. Note that the reception date should read "25 February 1916" instead of the 1915 date which is printed; 1915 would be before the collaboration with de Haas which is referred to in the opening paragraph. Prof. Dr. Treder, "Einstein," (ref. 23) has reprinted a hitherto unpublished precis of this work.

[38]Einstein and de Haas, "Experimenteller Nachweis" (ref. 27), p. 153.

[39]Adrienne R. Weill-Bruschwicg, "Paul Langevin," DSB, VIII, pp. 8-14, on p. 11.

[40]Ibid., p. 12.

[41]For more discussion of these problems, especially Ehrenfest's role in the quantum controversy, see Klein, Paul Ehrenfest (ref. 25), esp. pp. 264ff. See also T. Kuhn, Black-Body Theory and the Quantum Discontinuity, 1894-1912 (Oxford: Oxford University Press, 1978), esp. pp. 210-220 (on specific heats) and chapter 10, "Planck's New Radiation Theory," esp. pp. 264ff. (on Planck's 1911 theory); A. Pais, "Einstein and the Quantum Theory," Rev. Mod. Phys., 51 (1979), esp. pp. 878-883.

[42]Planck's zero-point energy outlived the theory from which it was derived. In quantum mechanics the zero-point energy of a harmonic oscillator may be thought of as a consequence of the uncertainty relation: let

$$E = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2$$

give the energy of a particle oscillating in a harmonic potential where m = mass, p = momentum, ω = frequency, and x = position. This energy E will not take its minimum at $p = 0$ since the uncertainty relation would then require the mean value of x to be infinitely large. If we let the mean value of $x = h/p$ and minimize E by setting $dE/dp = 0$, we get:

$$p^2/2m = h\omega/2.$$

Therefore the mean kinetic energy which minimizes the total energy of a quantum mechanical harmonic oscillator is $\frac{h\omega}{2}$. A more rigorous derivation requires the introduction of raising and lowering operators.

[43] A. Einstein and O. Stern, "Einige Argumente für die Annahame einer molekularen Agitation beim absoluten Nullpunkt," Annalen der Physik, 40 (1913), pp. 551-560. I would like to thank A. Pais for drawing my attention to this article and for a very helpful discussion.

[44] *Ibid*., p. 560.

The approach taken by Einstein and Stern seems unfamiliar to us now as we are used to computing the average rotational energy using quantum mechanical ideas:

$$\bar{E} = \sum_n E_n \exp\,[-E_n/kT] = \sum_j \frac{J(J+1)}{2A} \exp\,[-J(J+1)/kT],$$

where $J(J + 1)$ are the eigenvalues of the total angular momentum operator and A is the moment of inertia of the molecule. For low temperatures, only the first terms contribute and one obtains an expansion which goes to zero as the temperature goes to zero. This is also true of the specific heat. Physically, this is simply saying that the lowest energy state of rotational motion is no rotation at all—there is no zero-point energy associated with rotational angular momentum. This anachronistic aside, however, has little to do with the intentions of Einstein and Stern in 1913.

[45] *Ibid.*, p. 560.

[46] Klein, *Paul Ehrenfest* (ref. 25), pp. 256ff.

[47] Einstein to Besso, letter of 12.2.15 from Berlin. In *Albert Einstein - Michele Besso Correspondence 1903-1955* (Paris: Hermann, 1972), pp. 57-58.

[48] Niels Bohr, "On the Constitution of Atoms and Molecules," *Phil. Mag.*, *26* (1913), p. 1.

[49] Letter from Hevesy to Rutherford 14 October 1913 cited in Klein, *Paul Ehrenfest* (ref. 25), p. 278.

[50] G. Holton, *Thematic Origins of Scientific Thought* (Cambridge: Harvard University Press, 1973), esp. pp. 362-367.

[51] M. J. Klein, "Einstein and the Wave-Particle Duality," *The Natural Philosopher*, *3* (1964), pp. 5-49, on p. 7.

[52] A. Einstein, "Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung," *Physikalische Zeitschrift*, *10* (1909), pp. 817-826.

[53] Einstein and de Haas, "Experimenteller Nachweis" (ref. 27), p. 152.

[54] *Ibid.*

[55] Einstein, "Lorentz" (ref. 25), p. 6.

[56] *Ibid.*

[57] John Perry, *Spinning Tops*. The "Operatives' lecture" of the British Association meeting at Leeds 6 December 1890 (New York: E. & J.

Young and Co., 1890). Reprinted as Spinning Tops and Gyroscopic Motion (New York: Dover, 1957), on p. 65.

[58] Ibid.

[59] Barnett, "Magnetization by Angular Acceleration," Science, 30 (1909), p. 413.

[60] Ibid.

[61] A. Schuster, "A Critical Examination of the Possible Causes of Terrestrial Magnetism," Physical Society of London, Proceedings, 24 (1911-12), pp. 121-137.

[62] Ibid., p. 122.

[63] Samuel Barnett, "Magnetization by Rotation," Physical Review, 6 (1915), pp. 171-172, on p. 171. [Hereafter referred to as (1915a).] A longer version of the article with the same title appears in the same volume, pp. 239-270, on p. 270. [Hereafter referred to as (1915b).]

[64] Schuster, "Terrestrial Magnetism," (ref. 61), p. 122.

[65] Barnett (1915b) (ref. 63), p. 269.

[66] Barnett, "The Magnetization of Iron, Nickel, and Cobalt by Rotation and the Nature of the Magnetic Molecule," Physical Review, 10 (1917), pp. 7-21, on p. 7.

[67] Ibid., p. 8.

[68] Ibid. Emphasis added.

[69] Barnett (1915b) (ref. 63), p. 255.

[70] Ibid.

[71] Barnett (1917) (ref. 66), p. 21.

[72] I. Q. Stewart, "On the Moment of Momentum Accompanying Magnetic Moment in Iron and Nickel," Physical Review, 11 (1918), pp. 100-120.

[73] Ibid., p. 120.

[74] Ibid. Stewart must mean "angular momentum" since if one equates the angular velocities, the g-factor cannot possibly be 2. I thank E. M. Purcell for this observation.

[75] Emil Beck, "Zum experimentellen Nachweis der Ampèreschen Molekularströme," Annalen der Physik, 18 (1919), pp. 109-148 on p. 113.

[76] Ibid.

[77] Ibid., pp. 122-125.

[78] Ibid., p. 144.

[79] Unpublished letter from Einstein to de Haas 9 September 1919. See Ref. 36.

[80] G. Arvidsson, "Eine Untersuchung über die Ampèreschen Molekularströme nach der Methode von A. Einstein und W. J. de Haas," Physikalische Zeitschrift, 21 (1920), pp. 88-91.

[81] Ibid., p. 90.

[82] Institut Internationale de Physique Solvay, Atomes et Electrons: Rapports et Discussions du Conseil de Physique (Paris: Gauthiers-Villars, 1923), pp. 206-227.

[83]*Ibid.*, p. 214.

[84]Einstein and de Haas, "Experimenteller Nachweis" (ref. 27), p. 164.

[85]*Atomes et Electrons* (ref. 82), p. 226.

[86]Transcription of meeting at Reichsanstalt (March 1922) in Berlin, reproduced in Christa Kirsten and Hans-Jurgen Treder, *Albert Einstein in Berlin 1913-1933* (Berlin: Akademie Verlag, 1979), Volume II, p. 161.

[87]S. J. Barnett, "Further Experiments on Magnetization by Rotation," Proceedings of the Philosophical Society of Washington, *Journal of the Washington Academy of Science*, *11* (1912), p. 163.

[88]*Ibid*.

[89]S. J. Barnett, "Additional Experiments on the Nature of the Magnetic Molecule," *Physical Review*, *17* (1921), pp. 404-405.

[90]S. J. Barnett, "The Angular Momentum of the Elementary Magnet," *Bulletin of the National Research Council*, vol. 3, part 3, number 18 (1922), pp. 235-268.

[91]*Ibid.*, p. 242. Barnett cites M. Abraham, *Ann. d. Phys.*, *10* (1903), pp. 151, 169, 171.

[92]Unpublished letter from Bauer to J. C. Merriam (President of the Carnegie Institution), 28 November 1922. From the file marked "S. J. Barnett" in the archives of the Carnegie Institution, Washington D.C. I thank Dr. Franklin Portugal for his kind assistance.

[93]Letter from J. A. Fleming (Assistant Director) to L. A. Bauer, 28 November 1922. See ref. 92.

[94]Fleming to Bauer, 27 November 1922. See ref. 92.

[95]S. J. Barnett and L. J. H. Barnett, "New Researches on the Magnetization of Ferromagnetic Substances by Rotation and the Nature of the Elementary Magnet," Proceedings of the American Academy of Arts and Sciences, 60 (1925), pp. 126-216, on p. 215.

[96]Ibid., p. 128.

[97]Letter from G. Uhlenbeck to author 9 July 1979.

[98]A. P. Chattock and L. F. Bates, "On the Richardson Gyromagnetic Effect," Phil. Trans. A, 223 (1922), pp. 257-288. Cited in E. C. Stoner, Magnetism and Matter (London: Methuen & Co. Ltd., 1934), p. 239.

[99]C. N. Sucksmith and L. F. Bates, "On a Null Method of Measuring the Gyromagnetic Ratio," Proceedings of the Royal Society, 104 (1923), pp. 499-511.

[100]"Gyromagnetic Effects: History, Theory and Experiments," Physica, 13 (1933), pp. 256-268, on p. 266. By the time this article appeared, spin had become an accepted part of the theory of the electron. Barnett wrote, "One of the most important parts of [Barnett's] investigation is...that the magnetic element consists primarily of a Lorentz electron spinning on a diameter, and not an electron moving in an orbit." (Ibid., p. 254, emphasis in original.) To this the following is footnoted: "As suggested by O. W. Richardson, Inst. Internat. de Phys., Solvay 1921, also a little later by E. H. Kennard,...L. Parson

and others had suggested long before that the electron itself possessed a magnetic moment." There is no mention of Goudsmit and Uhlenbeck. It goes without saying that all of the authors Barnett mentioned were thinking of a classical spin (as Barnett himself still seems to be doing in 1933) by speaking of an electron's diameter.

[101] "Review of Gyromagnetic Ratio Experiments," Reviews of Modern Physics, 34 (1962), pp. 102-109. For a review of detailed quantum mechanical calculations of the Einstein-de Haas effect see S. P. Heims and E. T. Jaynes, "Theory of Gyromagnetic Effects and Some Related Phenomena," Reviews of Modern Physics, 34 (1962), pp. 143-165.

[102] T. S. Kuhn, "The Functions of Measurement in Modern Physical Science," Isis, 168 (1961), pp. 161-193.

[103] Ibid., p. 197.

[104] Ibid., pp. 200-201.

[105] Ibid., p. 194.

[106] Another account of the role of theoretical dispositions is given by Norwood Russell Hanson, Patterns of Discovery (Cambridge: Cambridge University Press, 1975) (esp. ch. 1). According to Hanson, our theoretical biases put us in a frame of mind similar to the one we are in when we see a duck rather than a rabbit in a Gestalt picture. Such an explanation, however, requires a great leap of faith to believe that the psychological processes relevant to a short-term psychological test are relevant to ten years of careful experimental work. It is very difficult to imagine an experimentalist of the caliber of de Haas simply

being mentally incapable of absorbing the message behind relevant meter readings for such a long time.

[107]S. J. Barnett and L. J. H. Barnett, "Improved Experiments on Magnetization by Rotation," Physical Review, 20 (1922), pp. 90-1, on p. 90.

[108]M. Deutsch, "Evidence and Inference in Nuclear Research," Daedalus (Fall 1958), pp. 88-98 on pp. 97-8.

CHAPTER III

THE DISCOVERY OF THE MUON AND THE RESURRECTION OF QUANTUM ELECTRODYNAMICS

When was the muon discovered? According to G. Bernardini, "the mu meson as a peculiar ionizing fraction of the bulk of cosmic rays was revealed by an experiment by Bothe and Kolhörster in 1929."[1] John Wheeler, however, considered that the theoretical work of Niels Bohr and E. J. Williams, together with the experimental work of C. D. Anderson and S. Neddermeyer, "established the existence of the muon" in 1936.[2] By contrast, Bruno Rossi spoke of "the discovery of the μ-meson in 1937" by Anderson and Neddermeyer and J. C. Street and E. C. Stevenson."[3] Street himself credited J. F. Carlson and Robert Oppenheimer in 1937 as first arguing for the necessity of the existence of a new particle of mass intermediate between that of the proton and the electron.[4] Finally, A. Pais began an article on the development of particle physics by stating that the mu meson was discovered by C. F. Powell in 1947.[5]

The discrepancy of over eighteen years in these dates indicates some of the problems implicit in looking back at past experiments. It is often impossible to determine in retrospect which experiment "demonstrated" the existence of a new particle or effect. A more interesting question is what kind of evidence _at the time_ convinced the experimentalists that they were looking at a real effect and not at an artifact of the machines or the environment. In other words, we want to know on what grounds the experimentalists decided to _end_ their experiments.

In the case of the muon we need to consider two very different, and often competing lines of thought on the nature of cosmic ray research in the early 1930s. One line was pursued by R. A. Millikan, Anderson, and some of their colleagues and students, principally at Cal Tech. The other line was developed on the theoretical side by H. Bethe and others, and on the experimental side, inter alia, by Rossi, Street, and A. H. Compton. Only by understanding the experimental and theoretical framework within which the two groups were working can we understand how Anderson and Street eventually came to believe that a new particle needed to be admitted to physics, and, at the same time, rescued quantum electrodynamics from one of its first serious crises.

1. Millikan, Cosmic Rays, and the Birth Cry Theory

One of the first indications that a penetrating radiation continually bombarded the earth came as early as 1903. In that year it was discovered that an electroscope discharged in an air-tight chamber at a slower rate if the thickness of the walls was increased.[6] This implied that an exterior source of the discharge must exist, though the nature of this cause was entirely unknown.

Then, in the 1910s, in one of the more adventurous espisodes in physics, A. Gockel, V. Hess, and W. Kolhorster took electroscopes with them on extremely high-altitude balloon flights to measure the rate of discharge at different heights.[7] They found that the rate increased with height and drew the conclusion that the "rays" causing the ionic discharge of the electroscope were entering the earth's atmosphere from above. Millikan, intrigued both by these experiments and by his own high-altitude meteorological work during the First World War, tried during 1921-23 to find the balloons necessary to explore the problem using self-recording electroscopes.[8]

The results of Millikan's work first appeared in print in a series of three articles in the Physical Review. The first, received 24 December 1925, reported that the radiation was considerably less absorbable than the Hess experiments had indicated.[9] By absorbability, Millikan and the other workers at that time meant the probability of a ray being absorbed per meter of water. This, in itself, was important, for in this and all other early cosmic ray experiments, it always tacitly was assumed that only the mass of the absorber needed to be taken into account. Still, Millikan and I. S. Bowen concurred completely with the earlier conclusion that the radiation had an origin outside of the earth's atmosphere.

In the second installment of the series,[10] Millikan, this time collaborating with R. M. Otis, conducted a detailed set of airplane and mountain peak electroscope measurements, hoping to eliminate spurious variations in the electroscope's discharge rates by carefully compensating for the systematic variation of temperature and pressure. Their procedure in recording the data was to place varying numbers of lead sheets about 5 cm thick each on top of the electroscope and to measure the rate of discharge. For their various high-altitude measurements, they again found a lower leak rate than Hess and Kolhörster had recorded. This led Millikan to conclude, "cosmic rays must be considerably more penetrating than Hess and Kolhörster had proposed (thus not producing many ions in the electroscope) or else the rays were of merely local origin with energies of ThD for instance."[11]

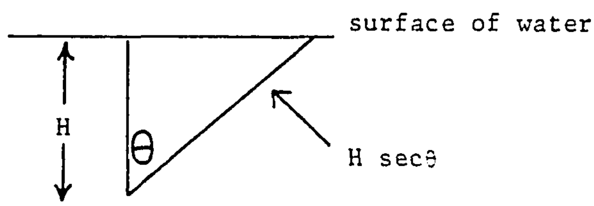
The theoretical presuppositions in both of these articles were always: (1) that the rays behaved like γ-rays and (2) that their absorption was a function only of the amount of mass (of air, lead, or water) they had to traverse. Finally, in the third and final part of the series, Millikan and G. H. Cameron made these more or less implicit assumptions explicit, setting out a theoretical account of the origin of cosmic rays.[12] This was a task that would preoccupy Millikan for the rest of his life.

Before the authors could be sure of the origin of cosmic rays, however, they needed to extend the results obtained on the mountain peaks. In the previous set of measurements, they had deduced that cosmic rays must be either too soft to penetrate the thin lead sheets or else so exceedingly penetrating that they failed to ionize the electroscope gas. On the latter supposition, they now sought an absorber thick enough to manifest the absorbability of the rays. The absorber most suited for this purpose was the water of a deep snow-fed lake (to avoid radioactive contamination) into which they could submerge an electroscope.

Moreover, by conducting their experiments in two different lakes at quite different altitudes, they could evaluate the effect of the intervening air mass. From these measurements they found that the two ionization rate vs. depth curves were precisely the same, except that measurements in the lower lake corresponded to deeper measurements of the upper lake. They took this to be good evidence that the rays did not originate in the intervening atmosphere, which acted only as an absorber.

The two physicists then calculated an absorption constant, μ, by which the rate of ionization should fall off as the cosmic rays penetrated the air and water. Assuming an isotropic intensity I_0 entering the atmosphere from above, they wrote an expression for the infinitesimal difference in flux as a function of solid angle and depth in the water:

$$dI = 2\pi I_0 \sin\theta \, d\theta \, \exp(-\mu H \sec\theta)$$



From this differential equation, and the measured quantities I and I_0, they could determine μ. They found that ". . . <u>no one coefficient</u> [of absorption] <u>would fit the whole curve</u>. . . . In other words, the radiation is not homogeneous but consists of a spectrum of wave lengths."[13]

In passing from the coefficient of absorption to the wavelength of the cosmic ray photons, Millikan had to use the only theory of absorption then available. This could be represented schematically as:

(absorption) = (compton scattering) + (photoelectric emission),

where both the compton scattering and photoelectric emission were functions of the incoming photon momentum.

From the broad limits on the absorption coefficients, Millikan obtained limits on the incoming frequencies of the photons, and with this data began to speculate on the origin of the cosmic rays. Eddington and Jeans[14] had earlier proposed that protons and electrons might annihilate each other in the stars, producing very hard γ-rays which would radiate outwards.[15] Millikan and Cameron objected to this because the rays would be too hard, instead proposing that nuclear changes were "going on not in the stars but in the nebulous matter in space, i.e., throughout the depths of the universe."[16] The changes they had in mind would be either "(1) the capture of an electron by the nucleus of a light atom, (2) the formation of helium out of hydrogen, or (3) some new type of nuclear change, such as the condensation of radiation into atoms."[17] All of these hypotheses had in common the formation of more organized states of matter out of more chaotic forms. Though in this paper the idea is still relatively undeveloped, it is the first printed statement of what Millikan soon called his atom-building hypothesis, or the "Birth Cry of atoms."

The Millikan and Cameron paper was the first realization of a theme Millikan would pursue vigorously for many years. Why was this theory, which almost no one accepted outside his immediate circle of students and colleagues, so attractive to Millikan? Several factors were undoubtedly at work. Robert Kargon has shown how Millikan's earlier interest in the structure of the nucleus and the transformation of elements played a role in the development of the Birth Cry theory.[18] Robert Seidel argued that during the 1920s and early 1930s, Millikan's dramatic claims served as an enticement to the foundations in their support of cosmic ray research.[19]

Yet another factor that must be taken into account regards the connection between Millikan's religious views and his theory of the origin of the elements. Like many American physicists of the time, such as Arthur Compton, Henry Rowland, and Edwin Kemble, Millikan was the son of a Protestant minister. Indeed, Millikan considered his religious upbringing crucial for his later life and for the remainder of his days the attempt to reconcile God and science was a recurring theme in his writing. In his words, "The first fact which seems to me altogether obvious and undisputed by thoughtful men is that there is actually no conflict whatever between science and religion when each is correctly understood."[20]

In general terms, Millikan was a staunch defender of what he called "essential" religious belief: a Universalist who espoused a religious outlook unfettered by the dogmas of fundamentalism, but at the same time preserved from materialistic atheism.[21] In more specific ways, though, Millikan tried to go further than simply to restrict the two belief systems of religion and science to non-overlapping domains. Instead, he hoped to exhibit the mutual dependence of the two systems.

Typical of Millikan's pronouncements on these matters is an excerpt from his book, Science and Life, where he asserted: "There have been two great influences in the history of the world which have made goodness the outstanding characteristic in the conception of God. The first influence was Jesus of Nazareth; the second influence has been growth of modern science, and particularly the growth of the theory of evolution."[22] "Evolution" consequently occupied a singular place in

Millikan's thought. By this term, Millikan understood not just the variation and selection of Darwinian evolution, but the more general concept of progress in the world, as applicable to the inorganic as to the organic world. Thus, while Darwin was elevated to an exalted role as the discoverer of change in species, Röntgen, Curie, and Becquerel were credited with the discovery of the evolution of the elements.[23]

For Millikan, the importance of the discovery of radioactive decay was that it strongly suggested that the inverse process was occurring somewhere in the universe. In 1921, Millikan wrote: "[W]ith radium and with uranium we do not see anything but decay. And yet somewhere, somehow, it is almost certain that these elements must be continually forming. They are probably being put together now somewhere in the laboratories of the stars."[24] For, as he repeatedly argued, just as God intervenes in the process of the evolution of animals, so he does too in the evolution of the elements. Together, inorganic and organic evolution help usher in the highest stage in religious thought, for they contribute to the "conception of progress [which] has entered the world."[25] Progress, according to Millikan, was the chief "contribution of science to religion, and a powerful extension or modification of the idea that Jesus had seen so clearly and preached so persistently. He had felt that benevolence and then preached it as a duty among men. Modern science has brought forward evidence for its belief."[26]

This theme of the reconciliation of religion and science permeated many of the books and articles Millikan wrote on non-scientific subjects. In 1923, Millikan circulated a petition among scientists, politicians, and Protestant religious leaders, in which the signatories testified to a belief in a "sublime conception of God which is furnished by science, . . . as wholly consistent with the highest ideals of religion . . . ."[27]

All of these influences, intellectual, practical and religious, played a role in the shaping and later in the defense of the Birth Cry theory. The theory fit in with his earlier research interests in the structure of the nucleus, it served as a dramatic rallying cry in his search for expanded physics research facilities, and it confirmed his earlier suspicion that somewhere higher forms of matter were continually coalescing to keep the universe from running down. Millikan's theory was his contribution to what he considered one of the most important issues facing science: it was the capstone on the theory of evolution, the greatest "contribution of science to religion."

2. New Support for the Birth Cry Theory

Once Millikan and Cameron "established" the broad limits of the energy spectrum of the cosmic rays, they turned in 1927 and 1928 to a modification of their quantitative methods. Millikan likened this stage of the research to his earlier refinements of his determination of the charge of the electron in his oil-drop experiments.[28] But in which aspect of the oil-drop experiments was Millikan seeking the analogy? In part, Millikan certainly was referring to the greater accuracy of each successive experiment. But the analogy to the oil-drop experiment probably goes deeper; for this reason, it is worth reviewing a few key aspects of Millikan's earlier work.

Gerald Holton has shown in his "Subelectrons, Presuppositions, and the Millikan-Ehrenhaft Dispute,"[29] that Millikan's oil-drop experiment for years involved him in a dispute not only over the discreteness of the electric charge, but over methodological and philosophical problems. Millikan supported atomic theory and in general a granular representation of nature. By contrast, Felix Ehrenhaft, then an Associate Professor at the University of Vienna, became increasingly antagonistic to atomic theory.[30] Along with this division lay another: Millikan spoke of "seeing" the electrons, his approach was pragmatic and his philosophy one of realism. Ehrenhaft, following Mach and Lampa, left the "reality" of electrons aside, preferring hypotheses to be judged by their predictive value alone.

Millikan felt his realistic, pragmatic approach to physics to be superior to the European, Machist attitude towards "mere hypotheses."

dence from a suggestive hint to a rationally ordered argument, the detail is essential. But in an effort at least partially to recover some of the longer larger historical changes, I have chosen three clusters of experiments with the approximate dates of 1915, 1937, and 1973.

By a "rationally ordered argument" I have in mind the presentation of experiments in textbooks or review articles where the conditions of the experiment are explained, the results presented, and their interpretation given. In general such presentations will be formulated as support for a theory. The task of the historian is often quite different. Instead of trying to reorder the conclusions in the most logically consistent fashion, the historian is more likely to seek the original conceptual framework under which the experiment was designed and the results understood. Even so, many historical accounts of twentieth-century experiments gloss over the ground between the design of the experiment and the interpretation of the result. In part, this may be because the primary attention of historians of twentieth-century physics has been focused on the development of theory: quantum theory and special relativity. When the focus is on the theorists, it seems that examination of the experimental work recedes into the background where it can be described in terms of designs and results.[2]

By contrast, a specific goal of these essays is precisely to explore the stage of experimentation between the experimentalist's original intent and his final conclusions. In this often hidden realm, "Whatever else the controversy [with Ehrenhaft] was about," Holton concluded, "it was also about two ancient sets of thematically antithetical positions: the concept of atomism and of the continuum as basic explanatory tools in electrical phenomena, and the use of methodological pragmatism versus an ideological phenomenology."[31]

Millikan's methodological presuppositions had direct consequences for his experimental physics. By the strength of his conviction, he set aside many measurements not in accord with his atomistic hypothesis. Others, working under different presuppositions, might have regarded these measurements as providing reliable data. The point to be made about this is not that Millikan was a bad scientist; quite the contrary, to have done otherwise would have meant that like Ehrenhaft he would have been confronted with an undifferentiated mass of mostly invalid data with every conceivable value of e.

Given this background, it is hardly surprising that Millikan was very intrigued with the search for the discrete bands of energy in the cosmic rays. His remark that he was now approaching a stage of cosmic-ray work comparable with the later oil-drop experiments may well refer to a stage of experimentation where he expected the discreteness of the bands to emerge from the background just as the atomistic charge had become clear seventeen years earlier. Moreover, in Millikan's early methodological precepts we may find the clue for his support of an easily visualizable, intuitive model of the nucleus, as well as his antipathy towards the highly abstract and seemingly idealistic theory of quantum mechanics with its wave functions and non-commuting algebra. Such considerations must have seemed a long way from the careful and pragmatic program Millikan had set out for himself and his colleagues.

With their new electroscopes, Millikan and Cameron measured the ionization rate as a function of depth in various lakes. This time, though, they were confident enough in the accuracy of their absorption curve to use it to determine a spectrum of absorption constants using the same theoretical model of absorption they had expounded in 1926. Millikan's new confidence in fitting the parameters to fit the absorption curve was, he claimed, justified by the smoothness of the ionization-depth readings.[32] From various pieces of the ionization curve, the authors then obtained absorption constants by fitting exponential curves to each segment of the curve at one meter intervals.[33]

On the basis of the chart reproduced below, they chose three coefficients to be representative of the three simple exponential absorption curves they claimed composed the observed curve. According to the authors, the fact that the "mean drops suddenly to 0.11," at 11 meters, "obviously means that the cosmic rays are not at all continuously distributed between $\mu = 0.2$ and $\mu = 0.7$."[34] This procedure was (to say the least) quite dubious, because there is no unique set of coefficients that will fit the curve. By judiciously choosing segments of this continuous absorption curve, one could find <u>any</u> average slope and therefore conclude that it was composed of any number of elementary exponential curves with <u>any</u> values of μ.

TABLE III. *Absorption coefficients at various depths, in meters of water, below top of atmosphere.*

Depth in meters of water beneath top of atmosphere	Absorption coefficient μ	Depth in meters of water beneath top of atmosphere	Absorption coefficient μ
8.45– 9.5	0.22	15–20	0.065
9.5 –10.5	0.20	20–30	0.057
10.5 –11.5	0.11	30–40	
11.5 –12.5	0.09	40–50	0.05
12.5 –15	0.07	50–60	

Figure 1. Depth below surface of atmosphere in "equivalent" meters of water versus best fitting absorption coefficient, assuming cosmic rays are photons.

Source: Millikan and Cameron, Physical Review 31 (1928), p. 927.

According to Millikan and Cameron, their results were analyzed and presented on 16 February 1928, "entirely without the guidance of any theory."[35] Only then, according to the authors,

> [A]fter we had made the foregoing empirical analysis, prepared the foregoing paper . . . and presented the results in detail to the physics seminar at the Norman Bridge Laboratory, our minds being up to this time completely unbiased by any knowledge as to whether bands might be expected, or if so where they might occur, we set at the task of seeing whether we could find any theoretical justification for their existence, or for their energy values.[36]

Within a few weeks, on 23 April 1928, Millikan and Cameron were prepared to offer this "theoretical justification."[37] Their theory was entirely in keeping with the speculation advanced in 1926--indeed, in harmony with sentiments Millikan had expressed in his non-scientific writing as early as 1921. Now, he wrote that the study of cosmic rays indicated "that the more stable and more abundant elements like helium (abundant in the heavens), oxygen, silicon and iron, are being formed at the present time out of the primordial positive and negative electrons . . . ."[38]

The theoretical justification was elaborated shortly afterwards in a major article in the Physical Review.[39] Taking Aston's measurements of the mass defect of various atoms (i.e., the difference between the mass of an atom and the mass of the number of hydrogen atoms supposed to constitute the atom), one could calculate the energy released by the formation of nuclei by using Einstein's equation, $E = mc^2$. Then, using $E = h\nu$, Millikan could calculate the energy of the photon released in such "atom-building" (here m is the mass defect, h is Planck's constant, and ν is the frequency of the light emitted). These photons constituted the primary cosmic rays; any other particles were due only to secondary production in the earth's atmosphere.

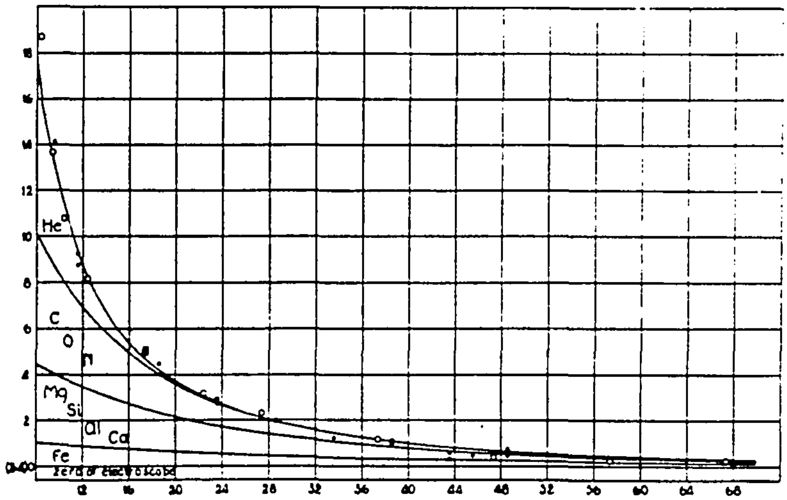
The "proof" of the process consisted, in 1928, of showing that (1) the photons of these specific energies would have just the absorption coefficients found in the ionization/depth experiments, (2) that the cosmic rays were not associated with the sun or other stars, and (3) that atomic disintegrations provided photons of too little penetrating power to explain experiments.

On the first point, Millikan and Cameron found an extraordinary agreement between theory and experiment.[40] For the production of oxygen, nitrogen, helium and silicon (these being the most abundant elements on earth added to the most abundant elements in space), the authors found the following correspondence of theory to experiment:

Atom-Building Process	Absorption Coefficients (Theory)	Absorption Coefficients (Experiment)
(oxygen and nitrogen produced by hydrogen)	.080	.08
(helium produced by hydrogen)	.30	.30
(silicon produced by hydrogen)	.041	.04
(iron produced by hydrogen)	.019	("not inconsistent with our data")

Figure 2. Atom-Building Processes. Theoretical *versus* experimental values for the absorption coefficient of gamma rays.

Using the experimentally-determined coefficients of absorption in conjunction with the then-current spectroscopic data for the relative presence of the elements outside the earth and its atmosphere, the authors plotted the absorption curve they would predict for the resultant cosmic ray photons, and compared it with experiment. Again, the agreement was extraordinary:



Comparison of experimental data with a built-up curve compounded from four absorption coefficients. Abscissas: depth in equivalent meters of water beneath surface of atmosphere; Ordinates: ionization in ions per cc per second. Dots are readings in Lake Arrow-head; circles, Gem Lake.

Figure 3. Comparison of experimental data with a built-up curve compounded from four absorption coefficients. Abscissas: depth in equivalent meters of water beneath surface of atmosphere; Ordinates: ionization in ions per cc per second. Dots are readings in Lake Arrowhead; circles, Gem Lake.

Source: Millikan and Cameron, Physical Review 32 (1928), p. 548.

With hindsight, we may note that by using the wrong particle (the photon) produced in a process which does not occur (atom-building of nitrogen, oxygen, etc.), and then using an absorption law which is also incorrect (ignores pair production, electron binding effects, etc.), Millikan and Cameron were able three times to produce a match between their theory and their experimental data (reduced in a quite arbitrary fashion) to one part in a hundred.

Millikan had a strong theoretical motivation for finding such an agreement. At stake were deeply-held religious views, his attempt to link his life's work to the religion of his upbringing. In an unusual conclusion to a scientific paper, Millikan and Cameron ended the Physical Review article with a remark about the role of God, reminiscent of beliefs Millikan had voiced at least several years before confronting the problem of the origin of cosmic rays. The target of Millikan's attack once again was the belief that God is in some way absent from the immediate world. Without atom-building, Millikan asserted, we are condemned to believe in the inexorable heat death of the universe. This heat death

> conflicts with no observed facts, and before the advent of Einstein it was a necessary consequence of the Second Law provided the universe were treated as a closed system. Scientists, however, have always objected that such treatment represents an extravagant and illegitimate extrapolation from our very limited mundane experience, and modern philosophers and theologians have also objected on the ground that it overthrows the doctrine of Immanence and requires a return to the middle-age assumption of a Deus ex Machina.[41]

Soon after the Millikan and Cameron paper appeared in print, Millikan received what must have been a rather distressing letter from

J. Robert Oppenheimer, then in Zurich. Oppenheimer, who undoubtedly viewed Millikan's fast and loose atomic physics with some scepticism, pointed out that the highly touted numerical confirmation of the band measurements was spurious. He wrote:

> Last year, when you were working on the interpretation of the absorption curves of the cosmic radiation, you asked me with what certainty the formulae of Dirac could be accepted. I answered, I think, that they could be taken as reliable, and that they could not be appreciably altered except by a fundamental change in the equations of physics. As you surely know, I was wrong to insist upon this reliability; the fundamental equations of the theory have in fact been altered; and there is a corresponding change for the absorption coefficient of hard radiation. The new formulae have been worked out by Klein and Nishina, and shewn, in the region in which you are interested, to give an absorption differing by as much as fifty percent from that calculated on the older basis.[42]

Oppenheimer went on to warn Millikan that even the new formula referred only to the scattering of light from free electrons and might very well require modification when one included nuclear effects. Once these nuclear effects were considered, he added, an extra-nuclear electron attached to a lead nucleus might behave very differently from an extra-nuclear electron associated with the nuclei found in air.

Oppenheimer's bad news hardly dissuaded Millikan from his theory. "I am," he replied to Oppenheimer,

> of course a little disappointed that the Dirac formula, which actually fits so well, has not the credentials which we thought it had a year ago. The quantitative fit, however, is only a part of the agreement, so that I do not think that the interpretations which I have given are as yet ready for the discard. There is no other interpretation that I can see for the sequence of frequencies which we observe quite independently of whether there is an exact quantitative fit as to the numerical values or not.[43]

Nonetheless, Millikan continued in 1930 to try to match the data to the Klein-Nishina formula.[44]

In print, Millikan reiterated the connection between God and cosmic rays in his retiring address as president of the American Association for the Advancement of Science on 29 December 1930.[45] Here Millikan outlined ten steps as "sign-posts on the road towards an answer" to the question of the "origin and destiny of all physical elements." First came the principle of conservation of energy; second, the second law of thermodynamics. This latter led some people to Millikan's nemesis, a "Deus ex Machina" who started a universe which would thereafter run down to its inevitable "heat-death."

Against this Deus ex Machina, science afforded the third sign-post, the theories of geological and biological evolution, which "identif[ied] the Creator with His universe, to strengthen the theological doctrine of immanence."[46] Fourth, physicists discovered the radioactive decay of certain elements; fifth, the great lifetime of the sun and stars; sixth, the interconvertibility of mass and energy. For Millikan, these last developments suggested the possibility of slow "age-long change" in the nature of the universe.

The seventh sign-post was the discovery that elements had the weights of integral multiples of hydrogen; eighth, that positive and negative electrons could annihilate in the nucleus creating "a terrific death-yell" of radiation which would heat the interiors of suns; ninth, that Aston calculated the mass defects of the various elements. These advances also contributed, in Millikan's opinion, to the cause of refuting the doctrine of the Deux ex Machina.

But capping this whole line of thought were the various cosmic ray measurements of Millikan, which were discussed above: cosmic radiation comes in bands, is isotropic, and is more penetrating than the hardest gamma ray. Thus, Millikan's tenth and crowning discovery provided the conclusion of an essentially theological path laid out by physics. There is "a bare suggestion" that

> If atom formation out of hydrogen is taking place all through space, as it seems to be doing, it may be that hydrogen is somehow replenished there, too, from the only form of energy that we know to be all the time leaking out from the stars to interstellar

> space, namely radiant energy. This has been speculatively suggested many times before, in order to allow the Creator to be continually on the job. Here is, perhaps, a little bit of experimental finger pointing in that direction.[47]

Millikan hoped to bolster his theory by the studies of one of his young post-doctoral students, Carl D. Anderson. Anderson had been both an undergraduate and graduate student at Cal Tech, completing his doctoral dissertation in 1927 on the space-distribution of photoelectrons using a cloud chamber. In 1930, Millikan suggested that Anderson undertake a study of the "energy of corpuscular radiation emitted by the incident photons," using the cloud-chamber techniques with which Anderson was already familiar.[48]

These experiments, Millikan hoped, would provide better data on the primary energy of the photons than could be gathered from the absorption experiments; for, as Millikan by then knew from Oppenheimer's letter, the relation between incident photon energy and absorption was not at all clear. It seems safe to assume that Millikan expected that the secondary electrons which Anderson would observe in the cloud-chamber experiments would exhibit the supposed band structure of the primary cosmic-ray photons.

The first paper published on these cloud-chamber experiments was a joint effort by Millikan and Anderson in 1932,[49] in which they found positive particles on cloud-chamber photographs. The authors interpreted these results as being evidence of nuclear disintegrations, identifying the positive particles immediately with the proton. In this way, they introduced a new process by which radiation could be absorbed by matter in addition to compton scattering and photoelectric emission. In part, Millikan's readiness to accept the role of nuclei in photo-absorption processes might be due to the comments by Oppenheimer in his letter. Still, Millikan and Anderson's ideas on the structure of the nucleus remained far behind the times and entirely without quantum mechanical ideas. Instead, as in his electron work two decades earlier, Millikan maintained a visualizable concept of the nucleus: electrons, positrons, and protons bound together and occasionally released by a sufficiently energetic photon. If Millikan was antipathetic to the new quantum mechanical studies of the nucleus, he remained quite convinced of his own atom-building theory. Concluding the joint paper, the authors reiterated Millikan's 1926 claim, adding to it the nuclear disintegration they had observed: "In a word, then, on the assumption that the tracks are due in all cases either to protons or to electrons, nine-tenths of all observed encounters yield energies which lie within the ranges computed from the Einstein equation and the atom-building hypothesis.[50]

Indeed, so convinced was Millikan that his theory was correct that he speculated that even the "one-tenth" of the "secondary" protons and electrons above 216 meV might well turn out to be only apparently so energetic, having been straightened out by turbulence in the cloud chamber.[51]

Anderson's photographs soon brought unexpected results. Among the cosmic ray cloud-chamber photographs he began to find positive particles. On 3 November 1931, he reported the findings to Millikan,[52] commenting that they show the "presence of positive particles as well as electrons indicating nuclear disintegrations by cosmic rays." The positive particles were thought to be either α particles or protons, and often to be simultaneously ejected from the nucleus with an electron. Finally, Anderson reported the "simultaneous ejection of three particles in at least one instance." Anderson concluded the letter by asserting that "A hundred questions concerning the details of these effects immediately come to mind . . . . It promises to be a fruitful field and no doubt much information of a very fundamental character will come out of it."

Anderson soon succeeded in obtaining clearer photographs by minimizing turbulence and improving the illumination in the chamber;[53] better measurements of curvature and ionization density therefore could be obtained. In this way, he calculated the mass of the positive particle to be the same as that of the electron. Anderson's results were first tentatively published in a short letter to Science,[54] but the first full report was made in 1933 in Physical Review.[55] In the longer article, Anderson followed Millikan in the pre-quantum mechanical belief that positive electrons are located in the nucleus. These nuclear positive electrons, Anderson and Millikan assumed, were being ejected by incoming cosmic ray photons.

As could be expected, Millikan immediately made use of Anderson's cloud-chamber work to bolster his Birth Cry theory. If the incoming rays were very energetic charged particles, Millikan reasoned, they would produce photons when they collided with nuclei. These photons should be found "shooting out equally in all directions from the bombarded nucleus in case it be assumed that [the charged particles] . . . can produce such high energy photons at all." It is unclear why the photon production should be isotropic; nonetheless, Millikan took the predominance of downward travelling positrons to indicate that primaries were cosmic ray photons.

Anderson, following Millikan's original assignment, continued to measure the energy distribution of the charged particles. These too were mustered in favor of Millikan's theory. In 1933 Millikan wrote:

> [T]he third and perhaps the most complete demonstration of this conclusion [that primary cosmic radiation is composed of γ-rays] is furnished by Carl Anderson's measurements of the energies of the actually observed cosmic ray particles; for these measurements show that the majority of these energies lie below 600 million volts, Dr. Anderson and I having published the estimate that not more than a tenth of these tracks reach appreciably above the billion (10^9) volt range.[56]

Little by little, Millikan was forced to concede the existence of electrons with energies too high to be secondaries from iron "atom-building," the highest energy "band" he had earlier discovered.

Within a few months, Anderson's new positron discovery was elevated to the position of one of the strongest arguments for the Birth Cry theory. Anderson had noted that the positive electrons constituted about one-half of the ionizing rays, and that almost fifteen percent of them started from the same center as the negatives. In Millikan's words, Anderson's work

> constitutes the strongest sort of evidence that both tracks arise immediately from the disintegration of an atom which has been hit by a primary ray of some sort, in other words that the immediate ionizing agents in the cosmic rays are themselves secondaries released from the atoms of the atmosphere by non-ionizing primaries.[57]

Anderson continued to work within the framework of Millikan's research program. After the discovery of the positron, he turned to an investigation of the effects produced in a cloud chamber by the gamma rays of ThC". This, it was hoped, would establish that the cosmic rays behaved like gamma rays, adding evidence to Millikan's assertion that the primary cosmic radiation consisted of photons. Anderson concluded that,

> Certain general conclusions can perhaps safely be drawn from the experimental data at hand. We shall consider the primary beam at sea level to consist in greater part of photons, a point of view held by Professor Millikan for several years, and now given additional support by the fact that hard gamma-rays of ThC" have been found to produce positrons as do the cosmic rays.[58]

A second conclusion followed: photons could interact not just with free electrons, but also with nuclei by the production of positive and negative nuclear electrons. Here, in the interaction of radiation and matter, Anderson began to stray from Millikan's theory for the first time, pointing out that Dirac's theory of the electron "perhaps" could account for the "general symmetry in occurrence between the positives and negatives."[59] Still, the Dirac process of pair production remained for Anderson a means for the production of the charged secondaries from primary photons.

On many fronts, then, Millikan's theory seemed to be triumphantly embracing the new discoveries, as well as the series of successes secured earlier. Yet another victory was claimed by Millikan when, in 1931, he concluded a series of balloon tests with self-recording electroscopes; the highest flights had climbed to 16km. "This measurement," Millikan asserted in a speech to a Paris audience in November 1931,

> shows . . . definitely and unambiguously that the ionization in a closed electroscope does not continue to rise exponentially in the upper atmosphere as it would if the incident rays were currents of high energy a) negative electrons, b) positive electrons, or c) γ-rays in equilibrium with their secondaries. On the contrary, the absorption coefficient passes through a maximum between 9 and 16 kilometers, and then falls to lower values.

> This is precisely what we would expect if the rays penetrating in the atmosphere were γ-rays which necessarily penetrated to a certain depth in the atmosphere before coming to equilibrium with their secondaries.[60]

As Millikan would later concede, this maximum in the upper atmosphere simply does not exist. Simultaneously, Millikan failed to find an effect that does exist, a phenomenon which has come to be known as the "latitude effect." Around the earth is a magnetic field which extends far beyond the atmosphere into space. As a result, if the primary cosmic ray particles are charged, near the poles there should be a higher flux of particles than at lower latitudes. This variation became widely known as the "latitude effect," and Millikan and his colleagues made frequent attempts to test it--for instance, by comparing flux rates in Churchill, Manitoba and Pasadena, California. By 1931, they had found "not a shred of evidence" for the latitude effect, and Millikan once again celebrated the success of the atom-building hypothesis.[61]

Like the ionization maximum that Millikan found in the upper atmosphere, the absence of a latitude effect provided yet another piece of supporting evidence for his atom-building hypothesis. Moreover, unlike his earlier speculations on bands, this time Millikan was making a claim which bore directly on experimental evidence. The gauntlet was thrown down. Arthur Holly Compton picked it up.

3. The Collapse of the Birth Cry Theory

Starting from an amicable relationship, Millikan and Compton entered into one of the most acrimonious and publicized scientific disputes of the century. It is evident from the Millikan archives that both men became very personally involved, even casting aspersions on each other's scientific integrity. Obviously amused at the spectacle of two Nobel laureates engaged in such a "dogfight" (as Millikan sometimes referred to it), the press raised the issue to front-page news.[62] Paralleling Millikan and his colleagues' studies of the latitude effect, Compton collaborated with a group also conducting a large-scale geographical survey of cosmic ray intensity. They concluded, in a preliminary report to the Physical Review,[63] that the latitude effect indeed existed, and sufficiently strongly to exclude Millikan's argument that all the charged particles seen at sea-level were secondaries produced within the earth's atmosphere.

Meanwhile, Millikan, probably feeling somewhat under siege by the Compton results, began to emphasize that his theory too could admit some latitude effects, since the primary photons could knock off some electrons from interstellar matter. As he wrote to Compton in late November of 1932, "Without modifying in any way anything I have ever written I can admit the possibility of some equatorial latitude effects so that we can appear before the public as not having got contradictory experimental findings."[64] Nonetheless, as late as 1936, Millikan commented that Neher's results "make it look as though there were no latitude effect at all, or if any a very small one . . . ."[65] (Neher had been a PhD student of Millikan's at Cal Tech, and in 1936 was an instructor in physics.) Millikan continued to believe for some time that no latitude effect existed.

The challenge to Millikan's theory presented by Compton was not the only difficulty Millikan faced. In 1929, W. Bothe and W. Kolhörster conducted an experiment with results that were potentially devastating to the Birth Cry theory.[66] Instead of focusing attention on the rate of discharge of an electroscope at different locations, the authors hoped to discover the nature of the rays themselves. To do this, they made use of the newly-developed Geiger-Müller tube, which is essentially a large cylindrical capacitor composed of a conducting hollow cylinder with a wire along its axis. The tube and the wire are then held at a constant and high potential difference. When a charged particle passes through the gas, it ionizes a few of the atoms. The ions, responding to the potential gradient between the center wire and walls, then rapidly begin to migrate, ionizing some of the atoms they hit. As the ions cascade in this manner, a current flows, and the capacitor is discharged. The surge of current could then be exhibited, for example, by means of an electroscope.

Bothe and Kolhörster's plan was to use two such Geiger-Müller tubes, each attached to a separate electroscope, the two tubes separated by a block of lead. If the tubes discharged simultaneously more often than would be expected by random discharges, this would be evidence for the passage of a single charged particle through the intervening lead. The principal difficulty in drawing such a conclusion was the rather poor time resolution provided by the simple observation of the two electroscopes. Nonetheless, Bothe and Kolhörster were convinced by their data that they had found cosmic rays to be primarily charged particles.

Bruno Rossi, then working in Florence, became interested in this aspect of Bothe and Kolhörster's work, and began to search for a way to improve their experiment.[67] Rossi's ingenious contribution to technique was a vacuum tube circuit which would emit a pulse only when two or more other pulses were delivered to the circuit at the same time.[68] Such a device was precisely what was needed for the cosmic ray work; in modified form, the "coincidence circuit" has become one of the most widely used tools in experimental physics.

Rossi wired three Geiger-Müller tubes to a coincidence circuit in such a way that a charged particle following a vertical path would discharge all three tubes and therefore cause a recording device to register the event. By inserting varying quantities of lead, Rossi was able to reaffirm the Germans' results with considerably more certainty. Indeed, Rossi found particles penetrating over a meter of lead.[69]

Anderson and Millikan struck back quickly. In his papers, Rossi had noted that secondary particles occurred in conjunction with the passing of a primary particle; unlike the primaries, the secondaries had a mean penetration of about one centimeter. (In retrospect, this fact could have provided the clue to the identification of the primary particles with a new type of particle and the secondaries with the electron; but for a variety of reasons, this would take almost five more years.) Anderson and Millikan too had noted the occurrence of these showers of secondary particles, and with S. Neddermeyer (then a post-doctoral student) and W. Pickering (a graduate student) they used this fact in conjunction with an important experimental observation: the number of shower particles increases up to about one and one-half cm of lead. This is, in fact, the case; however, they went on to infer that the number of shower particles would increase indefinitely with the thickness of lead. Rossi's coincidences, the four authors concluded,

> cannot in general be due to the passage of one charged particle through both counters and the intervening lead, but must be due to some mechanism by which a photon can release successively along, or in the general neighborhood of, its path a number of different particles whose separate but practically simultaneous action on the two or more counters is responsible for the observed coincidences.[70]

To argue for this position, the authors reiterated the kind of pre-quantum mechanical view of the constitution of the nucleus Millikan and Anderson had invoked many times before:

> [T]he simplest interpretation of the nature of the interaction of cosmic rays with the nuclei of atoms, lies in the assumption that when a cosmic ray photon impinges upon a heavy nucleus, electrons of both signs are ejected from that

> nucleus and appear in the form of the positrons and negatrons . . . . The larger, and the, in general, uneven numbers of positrons and negatrons . . . seem difficult to reconcile with the Dirac theory, as interpreted by Blackett and Occhialini, of the creation of electron-pairs out of the incident photons, and point strongly to the existence of nuclear reactions of a type in which the nucleus plays a more active role than merely that of a catalyst.[71]

By espousing the Dirac theory, P. M. S. Blackett and G. P. S. Occhialini had challenged Millikan and Anderson's interpretation of photon-nuclear encounters. Not only did Blackett and Occhialini interpret their photographs of positrons as due to Dirac pair production, their experimental apparatus made fundamental use of counters and coincidence circuits.[72] Millikan, by contrast, had on principle avoided any use of counter coincidence circuits which he mistrusted.[73] (Anderson only used random triggers of his cloud chamber for this reason.)

The Anderson, Millikan, Neddermeyer, and Pickering challenge to the work of Rossi, Dirac, Blackett, and Occhialini in many ways represented the last stand of the Millikan theory, although Millikan himself would continue to hold to it with various modifications until his death. For within the next few months, the latitude effect, the very high energy electrons, the penetrating particles, and the non-existence of an ionization maximum in the upper atmosphere were all widely accepted.

In this paper, the authors reiterated one last time all the old themes Millikan had stressed for almost a decade--a non-quantum nucleus, nuclear electrons, and photons as the primary cosmic radiation. But the fortress was collapsing on many sides, and soon even Anderson would definitively break away from Millikan's theory of the Birth Cry of atoms.

4. Bethe, Quantum Electrodynamics, and Cosmic Rays

Millikan, his associates and students were not the only group interested in cosmic rays and related problems. In Europe some physicists were deeply engaged in a very different style and type of research, both in theory and experiment. In particular, relativistic quantum mechanics sought (inter alia) to treat the high energy behavior of charged particles. Among the participants were P. A. M. Dirac, W. Pauli, W. Heisenberg, Niels Bohr, M. Born, Hans Bethe, and many others especially at Göttingen and Copenhagen. Because the problems of relativistic quantum mechanics often found their experimental tests in the behavior of high speed electrons and light, cosmic ray experiments became one of the proving grounds for the new physics, the other, of course, being in spectroscopic data.

It is not possible to do justice to all the complex developments of quantum electrodynamics in the 1930s here; topics such as the theory of holes, the interpretation of the Klein-Gordon paradox, self-energy problems, vacuum polarization, and so on. Instead this section will follow only that part of the early development of quantum electrodynamics, especially in the work of Hans Bethe, that brought theory into contact with experimental cosmic ray physics.[74]

The problem of the passage of electrons through matter dominated Bethe's theoretical work from his very first contact with physics. In 1926 when Bethe approached A. Sommerfeld for a problem to work on, he was assigned the task of accounting for some anomalies in the electron diffraction in crystals.[75] To solve the problem, Sommerfeld suggested that Bethe look for analogies with the diffraction of x-rays by crystals. Sommerfeld's advice could not have been more helpful; the pursuit of analogies between electron and light scattering became a hallmark of Bethe's work for the next decade.

After the wave-mechanical exercise, Bethe turned to the more complete quantum mechanical analysis of the problem for his doctoral dissertation. For this project, Bethe again looked to the diffraction of x-rays by crystals for analogical guidance, especially to the work of Paul Ewald's treatment of x-ray scattering. When he had taken his degree, Bethe went to Frankfurt and then on to Stuttgart where Ewald was a professor of theoretical physics. There he began work on what he later considered his best work,[76] "Towards a Theory of the Passage of Fast Corpuscular Radiation through Matter."[77]

In this paper Bethe applied Born's approximation scheme to the Schrödinger Equation to study the effect of atomic electromagnetic potentials on the passing electrons. As in his pre-thesis and thesis research, once again Bethe followed through the close analogy between electron and light scattering from matter. The parallel is as follows:[78]

<u>electron scattering</u>	<u>light scattering</u>
<u>Elastic Scattering</u>:	<u>Coherent Scattering</u>:
a) change in electron's direction without significant change in speed	a) no significant change in wave-length
b) no change in excitation of atom	b) interference is important
c) interference is important	

Inelastic Scattering:

a) usually small change in direction

b) change in excitation of atom

c) interference not important

Incoherent Scattering:

a) change in wavelength

b) change in target (Raman or Compton Scattering)

c) interference not important

Bremstrahlung:

slowing of electron in atomic field with emission of radiation

Photoelectric Effect:

absorption of photon with excitation of electron from atom

Bethe applied time dependent perturbation theory and the Born approximation to get a general expression for a rate of excitation of atomic states to the nth level.

$$d\phi_n(q) = 8\pi \frac{\alpha^2}{q^4} \left(\frac{M}{m} z\right)^2 \sin\theta \, d\theta \, |\varepsilon_n|^2 \frac{K}{K'}$$

where M = mass atom

m = mass projectile

z = charge on projectile

K, K' = initial, final wave number of projectile

$$\varepsilon_n(q) = \int \left[z - \sum e^{iqr} \right] \psi_0 \overline{\Psi}_n(r_j) \prod_j d\tau_j$$

q = momentum transfer from electron to atom

θ = scattering angle of final electron momentum relative to initial electron momentum.

From the differential cross section in this general form, the specific formula for the hydrogen molecule could be found, and integrated, to give the total cross section and eventually the stopping formula for the projectile.

The quantum mechanical result Bethe derived was somewhat different from an earlier, classical formula of Bohr.

$$\text{Bethe:} \quad -\frac{dE}{dx} = mv\,\frac{dv}{dx} = \frac{4\pi e^4 z^2 N}{mv^2} \sum f_{n\ell} \ln\left[\frac{2mv^2}{A_{n\ell}}\right]$$

where

f_{nl} = sum of all oscillator coefficients from electron level labeled n,l to any other level

A_{nl} = average excitation energy in level n,l

$$\text{Bohr:} \quad -\frac{dE}{dx} = \frac{4\pi e^4 z^2 N}{mv^2} \ln\left[\frac{mv^3}{(.78)\,2\pi R e^2 z}\right]$$

That is, the Bethe and Bohr results differed by a logarithmic factor:

$$1.12\ hv/(2\pi e^2 z)$$

which Bethe ascribed to the lack of detailed energy conservation implied by the quantum energy levels of the atom combined with the classical treatment of the electron in the Bohr theory.[79a]

Soon after Bethe's work was submitted, in the fall of 1930, he went to Cambridge, England, where he began discussing the problem with Blackett. Blackett, by then very interested in cosmic rays, encouraged Bethe to calculate a range-energy relation for electrons for comparison with experiment.[79b]

In 1931, Bethe spent several months with Fermi in Rome where he extended his work (on the passage of fast electrons through matter) to a relativistic treatment.[80] He then returned to Munich where he wrote on other subjects, including his famous review article on the electron theory of metals. By 1932, Bethe had found a job in Tübingen where the Nazis were beginning to make their presence felt through marches and protests; after the racial laws of April 1933 excluded Jews from State jobs, Bethe decided to emigrate,[81] leaving for a job at the University of Manchester, England.

From Manchester, Bethe frequently went to Cambridge, where he participated in the physics seminars with P. M. S. Blackett, J. D. Cockcroft, R. Peierls, W. Heitler, and others.[82] It was here that Heitler presented his work on the stopping of fast particles in matter, making use for the first time of the cross section of the Dirac pair creation process.[83] Surprisingly, Heitler found the cross section to increase (logarithmically) with energy. After Heitler used this formula to calculate the energy loss per centimeter, he commented:

> The theory seems to be here in disagreement with experiment. On the other hand, perhaps one should not expect the theory to give correct results for energies greater than $137mc^2$, since the wave-length for them becomes smaller than the classical electron radius e^2/mc^2 and Dirac's wave equation probably no longer applies.[84]

Heitler felt that something else was wrong; the increasing cross section with energy seemed to indicate a deep problem. In a letter to Bohr, he wrote: "Naturally, this shows that for very large energies the theory becomes false."[85] After listening to the oral presentation of Heitler's work at the seminar, Bethe began to wonder if both the energy loss discrepancy with experiment and the increasing cross section might be accounted for by taking into consideration the screening effect inner electrons would have on the nuclear electromagnetic field.[86]

In late February, Bethe and Heitler submitted what has become a standard reference paper in which they calculated the energy loss of a charged particle as it passed through matter including the effects of pair creation.[87] Bethe and Heitler's first order calculation avoided discussion of higher order infinities, and produced a model for the first order, relativistically correct approximations that characterized quantum electrodynamics before its reformulation in the 1940s by Richard Feynman, Julian Schwinger, and Freeman Dyson. The authors reasoned as follows:

A particle of momentum p_0/c and energy E_0 makes a transition to a momentum p/c and energy E. A light quantum is emitted of frequency ν such that $h\nu = E_0 - E$. The perturbing potential consists of an electrostatic part $V = Ze^2/r$ and a magnetic part, $H = -e(\vec{\alpha}A)$, where $\vec{\alpha}$ is the Dirac matrix vector. There are, in all cases, two events involving the electron: a photon is emitted, and the electron interacts with the nuclear field. Between these two events the electron is in what Bethe called an intermediate state. There are two possible such states: Either 1) no light quantum is yet present in which case the electron's momentum is p' = p + k, or 2) a light quantum of momentum k is present and the electron's momentum is p" = p_0 - k. Then Fermi's golden rule gives the rate of the combined process as:[88]

$$W = \frac{2\pi}{\hbar} \rho_E \rho_K \left| \sum \frac{H_{EI} V_{IA}}{E' - E_0} + \sum \frac{V_{E II} H_{II A}}{E'' - E} \right|^2$$

where ρ_E, ρ_K are the density of final states.

A is the initial state of electron, momentum p_0; no light quantum present

E is the final state of electron, momentum p; light quantum present

I is intermediate state with no light quantum present

II is intermediate state with light quantum present.

When the matrix elements are inserted for a coulomb field, cross sections can be derived. The total cross section then is, for energies E_0, E, and K all large compared to mc^2:

$$\phi = \frac{z^2}{137}\left(\frac{e^2}{mc^2}\right)^2 \frac{dk}{k} \frac{4}{E_0^2}\left(E_0^2 + E^2 - \frac{2}{3}E_0 E\right)\left(\ln \frac{2E_0 E}{hmc^2} - \frac{1}{2}\right)$$

(A similar expression is obtained for pair creation in a nuclear field by a photon.)

With the cross sections, Bethe and Heitler could calculate the probable energy loss per cm.

$$\Delta E = \int N \phi(\nu)\, h\nu\, d\nu,$$

where N is the density of atoms per cubic centimeter. Before they did, they modified the cross section by taking into account the screened field: with screening the matrix elements V_{AI} and V_{EII} no longer contain simple coulomb fields. Indeed, when they had rewritten the modified field in terms of the atomic form factor,

$$F(q) = \int \rho(r) \exp[(i/\hbar c)(\vec{q}\vec{r})]\, d\tau$$

they obtained the expression,

$$V_{AI} = V_{EII} = \frac{4\pi\hbar^2 c^2}{q^2} [z - F(q)]^2$$

$\rho(r)$ in the above formula is the density of atomic electrons, q is the momentum transfer to the atom. From this result it followed that screening only became important for energies such that E, E_0, and $h\nu$ all were much greater than mc^2.

Screening became ineffective for $E_0 \ll 137mc^2Z^{-1/3}$ and became nearly total for $E_0 \gg 137mc^2Z^{-1/3}$. In the former case, Bethe and Heitler reproduced the earlier result of Heitler and Sauter; in the latter case, the cross section no longer depended on E_0 for a fixed ratio of $h\nu/E_0$:

$$\phi(\nu)\, d\nu = \frac{z^2}{137} \frac{r_0^2}{E_0^2} \frac{d\nu}{\nu} 4 \left[\left(E_0^2 + E^2 - \frac{2}{3} E_0 E\right) \log\left(137\, z^{-1/3}\right) + \frac{E_0 E}{9}\right]$$

Thus, one of the disturbing features of the old Heitler-Sauter formula was eliminated. Integrating the expression for the cross section in the complete screening, the energy loss expression was:

$$-dE/dx = \frac{N z^2 r_0^2}{137} E_0 \left(4 \ln 183\, z^{-1/3} + \frac{2}{9}\right) \quad \text{for } E_0 \gg 137\, mc^2 z^{-1/3}$$

Comparing the result with Bethe's earlier work, the authors could show the radiative and collision energy losses for lead became equal for electron energies of about 10 mev. In general, they found:

$$\frac{dE/dx\ (\text{radiative}}{dE/dx\ (\text{collisions})} = E_0Z/(1600mc^2).$$

Unfortunately, this made agreement with experiment worse. Indeed, the first sentence of the section on comparison of the theory with experiments was underlined, and read: "The theoretical energy loss by radiation for high initial energy is far too large to be in any way reconcilable with the experiments of Anderson."[89] By way of explanation, the authors offered the following:

> The de Broglie wave-length of an electron having an energy greater than $137mc^2$ is smaller than the classical radius of the electron, $r_0 = e^2/mc^2$. One should not expect that ordinary quantum mechanics which treats the electron as a point-charge could hold under these conditions. It is very interesting that the energy loss of fast electrons really proves this view and thus provides the first instance in which quantum mechanics apparently breaks down for a phenomenon outside the nucleus. We believe that the radiation of fast electrons will be one of the most direct tests for any quantum electrodynamics to be constructed.[90]

In October of 1934, an International Conference on Cosmic Rays was held in London. Present were the advocates of the atom-building hypothesis, Millikan, Bowen, and Neher, as well as Anderson and Neddermeyer who, although drifting from the conclusions of Millikan, continued to pursue the program of measuring cosmic ray absorption coefficients and energies. Also present was the author of the most complete quantum treatment of the physics of cosmic ray absorption, Hans Bethe, as well as the experimentalist who had most advanced the program that treated the cosmic radiation as particles, Bruno Rossi.

By the time of the London Conference, Anderson and Neddermeyer were comparing their experimental results not with Millikan's atom-building theory, but with the new quantum calculations. Shortly before the Conference, they submitted an abstract to the Physical Review, in which they asserted: "Measurements of the energies of secondary electrons produced in plates of lead and carbon by cosmic ray electrons . . . have shown that the distribution in energy of secondary negatrons is, within experimental certainty, in agreement with the distribution calculated from the theoretical cross section given by Carlson and Oppenheimer."[91]

Even this much agreement, however, did not last long. Within a few months, at the London Conference, Anderson and Neddermeyer hardened their conclusions, and by so doing helped precipitate one of the first of the many theoretical crises quantum electrodynamics would face in the years to come. They argued that, "While the [absorption] data presented above give evidence for the existence of rather large radiative losses, they constitute as well strong evidence for the breakdown of the theoretical formula in the energy range above 100 MeV."[92]

At the time this remark was written, only the Heitler-Sauter formula was available, which as mentioned earlier, was an approximation that left out the screening of the nucleus due to inner electrons. In a footnote added after the Conference, the authors noted that even the Bethe-Heitler theory (which did include screening) predicted radiative losses "too high to be reconciled with our experimental data, although the latter contain as yet too few cases where accurate measurements are possible, for a satisfactory comparison to be made."[93] The only explanation for the discrepancy seemed to the authors to be that the particles were protons and not electrons.

Such a hypothesis seemed to clash directly with two other results. First, if the secondary electron energy distribution was calculated on the assumption that the primary cosmic rays were composed of protons, it was found to be incompatible with the measured distribution. Second, if the primaries were protons, some of them should arrive at sea-level with low energy. And while at high energies, it was difficult to distinguish positive electrons from protons, at low energies it was easy. The absence of any low energy protons thus presented a further argument against the primary proton hypothesis.

"The above considerations," Anderson and Neddermeyer concluded,

> which are of a statistical nature, and necessarily subject to the gathering of further data, tend to favor the view that most of the high energy cosmic ray particles at sea-level have electronic mass. If further data prove this view to be correct then it is obvious that the present theory of radiative losses by electrons must be inapplicable in the range of very high energies.[94]

Despite their qualifications, the effect of Anderson and Neddermeyer's announcement added to the discouragement of many of the theorists. In retrospect, of course, we can say that they were getting their first (statistical) glimpse of the muon which we now know constitutes some 90% of sea-level cosmic ray charged particles. At the time, however, the existence of a new particle was not discussed. Instead, it was the theory, quantum electrodynamics, that was being put to the test (and failing) at high energies.[95]

Bethe, who attended the talk by Anderson and Neddermeyer, immediately conceded that the experiments boded ill for the Bethe-Heitler theory. In the discussion period after the talks, Bethe commented:

> The experiments of Anderson and Neddermeyer on the passage of cosmic-ray electrons through lead are extremely valuable for theoretical physics. They show that a large fraction of the energy loss by electrons in the energy range round 10^8 volts is due to emission of γ-radiation rather than to collisions, but still the radiative energy loss seems far smaller than that predicted by theory. Thus the quantum theory apparently goes wrong for energies of about 10^8 volts.[96]

Future experiments, Bethe concluded, would be necessary mainly to determine precisely at what energy the alleged breakdown became significant.

Word of Bethe's change of heart over the quantum theory was soon received in Germany, where Carl Friedrich von Weiszäcker, Werner

Heisenberg, and other physicists continued to do cosmic-ray related research. Von Weiszäcker wrote to Bethe in December of 1934, asking:

> Do you now actually believe in your radiation formula for $E > 137mc^2$ or not? Anderson's London report was not too clear to me on this point, but you have in fact spoken with Anderson himself. In the meantime, you published a note according to which it seems you now believe in the calculations [of the Bethe-Heitler theory], but I could not determine with any certainty whether the reversal was partial or total. Heisenberg mentioned to me that he wrote you that Weisskopf has found a hole-theoretical argument against the Fourier analysis.[97]

For Bethe the concession that quantum electrodynamics would break down was not easy. The Bethe Heitler theory was in his eyes a great success precisely by sidestepping the more "philosophical" objections that stemmed from less practical issues than energy loss measurements. Oppenheimer, in particular, even before the specific energy loss measurements, was worried about the theory. Since the discovery of the positron in 1932. many theorists had re-examined the Dirac theory, taking its solutions more seriously than previously. During the spring of 1933, Bohr lectured at Cal Tech, where Oppenheimer spent time with him discussing problems related to the problem of pair-creation.[98] After Bohr's lecture, Oppenheimer wrote that he was working on the theory of pair creation,[99] and again to his brother, Frank Oppenheimer, in

October, that he was still at it.[100]

By the fall of 1933, Oppenheimer was convinced that something was deeply wrong with the theory. "I think . . .," he wrote to George Uhlenbeck, "that the methods of the radiation theory give completely wrong results when applied to wave lengths of the order of the electron radius."[101] In fact, by June of 1934, Oppenheimer was on the verge of despair about the state of physics when he wrote his brother Frank: "As you undoubtedly know, theoretical physics--what with the haunting ghosts of neutrinos, the Copenhagen conviction, against all evidence, that cosmic rays are protons, Born's absolutely unquantizable field theory, the divergence difficulties with the positron, and the utter impossibility of making a rigorous calculation at all--is in a hell of a way."[102]

Oppenheimer was thus already deeply pessimistic about the state of quantum electrodynamics before all of Anderson's results were finished and presented at the London Conference. In November, when these last figures were available, Oppenheimer submitted an article to the Physical Review, entitled "Are the Formulae for the Absorption of High Energy Radiations Valid?"[103] His answer to the title question was a resounding "no." The problem, Oppenheimer stressed, was two-fold: first, Bethe-Heitler theory predicted an increase of cloud chamber ionization with energy that was not observed by Anderson and Neddermeyer, or by Kunze. Second, the specific energy losses measured by Anderson and Neddermeyer also seemed too low to be compatible with Bethe-Heitler theory. Consequently, Oppenheimer argued that, "It is . . . possible to do justice to the great penetration of the cosmic rays only by admitting that the formulae are wrong, or by postulating some other and less absorbable component of the rays to account for their penetration."[104]

Oppenheimer concluded that the formulae do indeed break down at high energies; he then sought to explain why this should occur. The argument he offered was based on a consideration of classical (Lorentzian) electron theory. Suppose that an electron is given as a spherically symmetric distribution of charge in a sphere of radius $\rho = e^2/mc^2$. Then the theory only will give the correct motion of the particle if the radiative reaction of the electron can be treated as a perturbation on the motion given by the electron's direct response to the external field. That is, the force on the electron must be expressible as:

$$F = m\ddot{x} + 2e^2\ddot{x} + O(e^2\rho\ddddot{x}/c^4)$$

This series converges only if the terms decrease sufficiently rapidly, i.e., if their successive ratios are small:

$$\dddot{x}\rho/(\ddot{x}c) \ll 1; \qquad \ddddot{x}\rho/(\dddot{x}c) \ll 1; \ldots \quad etc.$$

In the rest frame such a condition can be given as a simple restriction on the frequency of the external field: letting $\bar{\nu} \equiv c/\rho = mc^3/e^2$, the condition $\dddot{x}\rho/\ddot{x}c \ll 1$ imples $\dddot{x} \ll \bar{\nu}x$. Oppenheimer's argument was that _if_ high energy frequency components of the external field were present and significantly larger than would be expected from classical theory, then the expression for the force would not converge. The theory would break down. Using a normal Fourier decomposition of the field, such high frequency components would not appear to be a problem. However, according to some recently completed work of Born,[105] it seemed that a more accurate electrodynamics would require non-linear expressions, and so the high frequency component of the external field might turn out to be much larger than that given by the usual Fourier decomposition. In this case the theory would fail. Similar considerations may have been behind Weisskopf's "hole-theoretical argument against the Fourier analysis," referred to in the von Weiszäcker to Bethe letter cited above.

Oppenheimer's students, among them Wendell Furry, also adopted a highly skeptical stance towards the validity of quantum electrodynamics at high energies. After recalculating the number of pairs produced by photons of different energies, Furry and J. F. Carlson asserted that at low energies their results agreed well with experiments on beryllium (which produces γ-rays of about 5 MeV). However, "For energies above twenty million volts the predicted pair production is even greater than that computed by Oppenheimer and Plessett, and hence even more irreconcilable with experiment. It seems possible to connect this discrepancy with the fundamental inadequacies of quantum electrodynamics."[106]

In sum, the high frequency components of an external field put into question the validity of the new quantum electrodynamics if the field could not be Fourier decomposed in the usual way. And this was just the claim of Oppenheimer on the grounds of Born's non-linear electrodynamics, and of Weisskopf on the basis of some (now unknown) consideration of Dirac's hole theory. Other physicists, like Nordheim, Furry, Carlson, Bohr, and Bethe, were also for some time convinced that the quantum theory must break down at high energies because of the new experimental results. The choice was clear: quantum electrodynamics or a new less absorbable ionizing radiation. For the moment at least, the entire circle of physicists interested in these problems opted for the collapse of quantum electrodynamics. In a letter to E. J. Williams, Bohr wrote: "I am ever more and more inclined in the experimental results to see an indication of a new fundamental aspect of electron theory, for which the limitation of classical theory may well leave room, but for which it offers no guide whatsoever."[107] A radically new theory seemed to be needed.

5. The Resurrection of Quantum Electrodynamics and the Discovery of the Muon

Oppenheimer's choice (Bethe-Heitler theory or a new particle) was complicated by the uncertainty as to which tracks the Bethe-Heitler theory ought to apply. Such was the confusion during this time that Anderson and Neddermeyer began to speak, among themselves, of "red and green electrons," where the red ones were especially absorbable and the green ones passed easily through matter.[108] In addition, by far the most dramatic phenomena observed in the cloud chamber and counter devices were the showers. Almost immediately the question arose whether the constituent particles of the showers were electrons, some other type of electron, or a new particle altogether. Both Rossi and Street later recalled that during this time in 1935, it was commonly assumed that the shower particles were the "new" type of particle and that the penetrating particles were ordinary high energy electrons that simply did not obey Bethe-Heitler theory. Such a conclusion was natural given the seemingly strange behavior of showers in which it often seemed that many particles were ejected from a single site.[109]

During Street's last year at the Bartol Institute (1932-33), he had constructed logic circuits, based on Rossi's publications, in order to conduct experiments on these showers and related phenomena.[110] All of these experiments were based on varying the geometric arrangement of counters that were wired to coincidence and anti-coincidence circuits. Since the counters could be separated by large distances, they were easily adapted to measuring the particle flux coming from a certain direction. It was therefore relatively easy for Street to compare the charged particle flux arriving from different directions. If the incoming cosmic ray particles were charged, and predominantly of one type, the earth's magnetic field would create an asymmetry between the flux from the east and that from the west. Like the latitude effect, the so-called "East-West Effect" became a test of whether the primary particles were photons or charged corpuscles. Unlike the latitude effect, the sign of the incoming particles is revealed in the direction of the asymmetry.

For this reason, and because the East-West Effect was free of the problems associated with transporting equipment to climatically and geographically diverse points, from 1931 on Rossi's work focussed more on this effect rather than on the latitude effect.[111] In 1933, Rossi and Benedetti, Thomas Johnson, and Luis Alvarez and A. H. Compton had independently found the effect, indicating to their surprise that the incoming particles were positive. When Street confirmed their results in late 1933, it left him with the firm conviction that primary particles were charged and not, as Millikan had claimed, photons.[112]

After his experience with the East-West Effect and related measurements, by the time Street came to Harvard in the fall of 1933, he was convinced that the coincidence circuits with Geiger tubes were an effective way to study cosmic rays. Ionization chambers, by contrast, did not seem to him to be of great interest.[113] In addition, from other preliminary measurements of showers, again using counters, he knew that non-ionizing radiation could produce secondary ionizing radiation.

At Harvard, Street repeated these experiments with his students, E. C. Stevenson, and later, with a student from MIT, L. Fussell. They were able to improve the apparatus and to make an absolute calibration of the counters. With this equipment, the Harvard group repeated Rossi's counter experiments showing that coincidences occurred between counters separated even by large thicknesses of lead (tens of centimeters). Thus when the Anderson, Millikan, Neddermeyer, and Pickering paper appeared in 1934 attacking Rossi's assertion that these coincidences were due to the passage of single particles, Street's work was implicated as well. Indeed, much later, Street recalled Millikan's absolute conviction that "nothing was going through all this thick material," since this would have conflicted with his Birth Cry theory. "So," Street remembers, "we thought we better learn how to do cloud chambers."[114]

Building the chamber was not easy, but when the project was completed, several important technical innovations made their use considerably easier. All cloud chambers work on the same principle: when a particle passes through a gas it ionizes atoms in its path. If the gas is subsequently suddenly expanded, the temperature drops and the gas will condense around the ionized particles, leaving a visible track. Ordinarily the expansion was triggered at a random time; this was, for instance, the technique of Anderson et al. But C. T. R. Wilson had explored the use of counter-controlled expansions, thus vastly increasing the number of useful photographs. Moreover, by 1934, since Street and his group were familiar with the use of logic circuits and Geiger tubes, they immediately wanted to combine the chamber with their logic circuits. In so doing, they created a device they called a hodoscope: a chamber sandwiched between two counters connected through a coincidence circuit. With this device they were able to vindicate both their own and Rossi's work by showing conclusively that individual charged particles were passing through at least 45 cm of lead. In their words, "at least 90% of the coincident counts for such an arrangement are directly due to the passage of single electrons through the apparatus."[115]

Almost simultaneously with the above publication, Street and R. H. Woodward (another of Street's graduate students) submitted an abstract that appeared in the same issue of Physical Review.[116] Before this abstract was printed, it was well known that showers increased to a maximum in about 1.5 cm of lead. From this fact, some authors had incorrectly concluded that the shower particles themselves had a peak in their penetration length of about 1.5 cm. Street and Woodward showed that this was not the case; individual shower particles were simply absorbed exponentially. By so doing, for the first time the authors focussed attention on the properties of shower particles as distinct from the showers themselves. Again, the key to their success was in the use both of counters and of a cloud chamber.

Yet another of Street's articles, in collaboration with Woodward and Stevenson, appeared in the same volume of Physical Review.[117] This time the authors were able to use their apparatus definitively to rule out the conclusion of the Anderson, Millikan, Neddermeyer, and Pickering paper, that coincidences of Geiger tubes surrounding large thicknesses of lead were due only to shower effects. Furthermore, Street, Woodward, and Stevenson combined their new and better absorption curves based on studies of individual particles with the energy distribution of Anderson, Millikan, Neddermeyer, and Pickering. Taken together, the two results determined a specific energy loss curve given as a function of energy. Still, as before, they invariably referred to the penetrating particle as the electron.

Since the penetrating particles were thought to be electrons, Street continued to search systematically for differences in character between the shower and penetrating particles. By January 1936, Street and Stevenson had gathered enough photographs of showers to draw some conclusions about the nature of the shower particles.[118] In particular, the authors contrasted the probability of an "electron," taken at random from the cosmic rays producing a shower, to the probability of a "shower electron" producing a shower. While the former was but two in a thousand, the latter was almost twenty-five percent.[119]

These "electrons taken at random" were measured by Anderson and Neddermeyer. By June 1936, Anderson, Neddermeyer, Rossi, Stevenson, and Street all knew that the electrons of Bethe-Heitler theory could not be reconciled with experiment. Anderson and Neddermeyer by this time had withdrawn all qualification from this judgment, and wrote: "It is obvious that either the theory of absorption breaks down for energies greater than about 1000 MeV, or else that these high energy particles are not electrons."[120] On the one hand, the authors recognized the success of the Bethe-Heitler theory in explaining the existence of large showers with altitude, the "transition effect," and the strong dependence of shower particles on thickness and type of material traversed. On the other hand, there remained a "large fraction of the sea-level particles" that were more penetrating than the theory possibly could permit.

To discover what these particles were, Anderson and Neddermeyer continued to search for penetrating particles of low enough energy to determine if they were in fact protons. By now asking the question, "<u>what are the penetrating particles</u>?" (and assuming the shower particles were electrons), they had reversed the earlier orientation of asking, "<u>what are the shower particles</u>?" (assuming the penetrating particles were electrons). Because the theory now seemed to be reasonably successful at describing what at first had seemed a qualitatively new phenomenon (showers), the new physics increasingly seemed to be associated with what had previously seemed to be the "ordinary" phenomena, namely the penetrating particles. Unfortunately, the much sought-after photographs of penetrating particles of sufficiently low energy remained difficult to obtain. In sum, by mid-1936, both the East and West Coast groups had made a conceptual separation of shower and penetrating particles, and there was a growing suspicion that it was the penetrating particles that were problematic by not conforming to the Bethe-Heitler theory of electrons' passage through matter.

However, unlike the penetrating particles, the shower particles were not very easily compared with the Bethe-Heitler theory, since by their very nature they were tied up in the often immensely complicated features of the showers. It was to bridge this gap between experiment and theory that J. F. Carlson and Oppenheimer set out to present a model of showers using the calculated cross sections provided by Bethe and Heitler.

For some time discussion of showers had been phrased in terms of the sequence of elementary processes:

annihilation of electron-positron pair, Bremstrahlung $\longrightarrow$ photons $\longrightarrow$ creation of electron-positron pairs

Nonetheless, no one had undertaken a quantitative analysis. Thus, the authors decided that from this qualitative scheme, they

> should like to derive on the one hand a further argument for the qualitative validity of the theoretical formulae and on the other for the often repeated suggestion that many showers are built up by a long succession of simple elementary processes, and not by the simultaneous ejection of a huge number of particles in one elementary act.[121]

Earlier attempts to analyze showers in any quantitative fashion were simply iterated calculations of the probability of a photon creating an electron-positron pair or freeing an electron by photoionization multiplied by the probability of the electron radiating a photon, etc. Such a procedure rapidly became unwieldly for complicated showers; for this reason, Carlson and Oppenheimer hoped to reduce the problem to a calculation of general diffusion equations. As a first approximation, they presented the following argument. In both pair production and radiation, two rays are produced for each incident one.[122]

Both processes occur in approximately the same length of matter, which they call $t = 1$. Then after t such lengths, approximately 2^t particles will be present. Since ½ of these will be electrons, their collective loss per length dt will be the number of particles times the individual rate of loss per length, $\partial E/\partial t$:

$$dE = (\tfrac{1}{2})2^t(\partial E/\partial t)dt.$$

The shower will come to an end after losing all its initial energy E_0, i.e., when it has traversed a distance T such that:

$$\int_0^T dE = E_0.$$

If $\partial E/\partial t = \beta$ is approximately constant, then

$$dE = E_0 = (\tfrac{1}{2})\beta\int_0^T 2^t dt = (\tfrac{1}{2})\beta 2^T/\ln 2, \text{ or}$$

$$2E_0 \ln 2 = \beta 2^T.$$

Therefore (i) the shower's length T increases only logarithmically with E_0, (ii) the number of particles 2^T will increase approximately linearly with E_0, and (iii) for showers of approximately 30 particles, $T = \ln_2 30 \sim 5$. For lead, one interaction length is about (½) cm, so the maximum would be around 5/2 cm of lead, which is in good accord with the observed maximum.

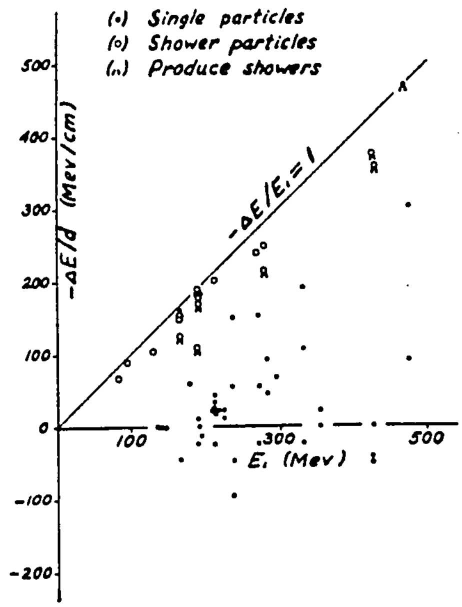
These (and the more precise predictions of the diffusion equations that followed) showed that the quantum electrodynamic view of showers as a compound of elementary processes could accurately represent many of the qualitative features of showers. Thus the showers seemed well accounted for. However, if the penetrating particles were electrons, according to these calculations they should be almost totally absorbed by 20 cm of lead; this was manifestly not the case. "From this," Carlson and Oppenheimer wrote,

> one can conclude, either that the theoretical estimates of the probability of these processes are inapplicable in the domain of cosmic-ray energies, or that the actual penetration of these rays has to be ascribed to the presence of a component other than electrons and photons. The second alternative is necessarily radical; for the cloud chamber and counter experiments show that particles with the same charge as the negative electron belong to the penetrating component of the radiation; and if these are not electrons, they are particles not previously known to physics.[123]

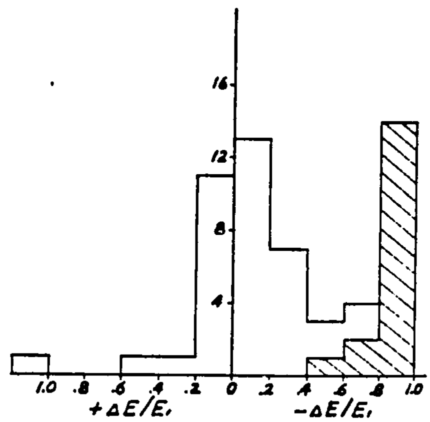
Indeed, since the success of the multiplicative shower theory, these particles "not previously known to physics" now were the main problem. As Oppenheimer understood, all his work on showers was valid "only if [one] admits the presence of another component [of cosmic radiation] to which the analysis is not at all applicable."[124]

The contrast between the two kinds of particles that had been pointed to by Street and Stevenson was then taken up again by Anderson and Neddermeyer in early 1937. Instead of looking at the specific energy loss of cosmic ray particles in general, they made separate measurements of shower and penetrating particles.[125] (See Figure 4 below.) To explain this separation between the two clusters of data points, the authors offered the following choice:

> Interpretations of the penetrating particles encounter very great difficulties, but at present appear to be limited to the following hypotheses: (a) that an electron (+ or -) can possess some property other than its charge and mass which is capable of accounting for the absence of numerous large radiative losses in a heavy element; or (b) that there exist particles of unit charge, but with a mass (which may not have a unique value) larger than that of a normal free electron and much smaller than that of a proton. This assumption would also account for the absence of numerous large radiative losses, as well as for the observed ionization. Inasmuch as charge and mass are the only parameters which characterize the electron in the quantum theory, assumption (b) seems to be the better working hypothesis.[126]



Energy loss in 1 cm of platinum.



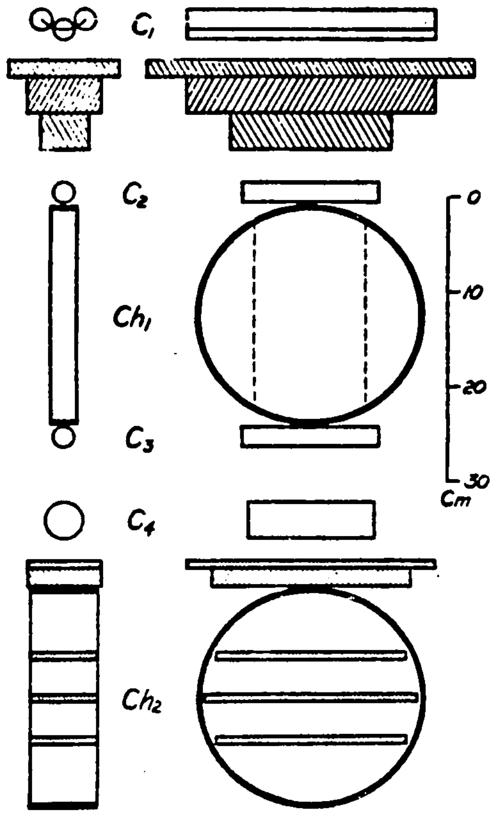
Distribution of fractional losses in 1 cm of platinum.

Figure 4. Source: Physical Review, 51 (1937), p. 884.[127]

Independently of Anderson and Neddermeyer's intermediate mass hypothesis, J. C. Street and his graduate students and colleagues, R. T. Young, Fussell, and Stevenson, arrived at the same conclusion. First, Fussell conducted studies of showers using a series of very thin (down to .07 cm) plates.[128] By so doing, the complex showers could be shown experimentally to follow the multiplicative pair production schema set out by Carlson and Oppenheimer. He concluded that his "observations give strong support to the radiation, pair formation theory of showers . . . ." It was then possible for Street to set up an apparatus to study the range, energy and shower production of the other, penetrating component of cosmic rays, assured that the Bethe-Heitler-Carlson-Oppenheimer theory accounted for the shower particles.[129] The new apparatus was built as indicated below (see Figure 5).

The first cloud chamber indicated whether the particle was single or part of a shower; the counter showed whether the particle continued through the apparatus; the second cloud chamber exhibited whether or not the particle produced showers. When the results were measured, it was clear that a much greater fraction (on the order of 10^4 more) particles were penetrating over 6 cm of lead than would be permitted either for electrons of similar momentum under the earlier measurements of Anderson and Neddermeyer or from the theory of Bethe and Heitler.[130] They too concluded there must be a new particle of intermediate mass.

Although it was the range-energy experiment that convinced Street and Stevenson[131] that a new particle existed, it is not for this work that they are usually remembered. In fact, most theorists were convinced by a remarkable photograph of a dense track, with an ionization



Apparatus for the study of the range, energy, and shower production of the penetrating particles. The upper cloud chamber is in a magnetic field of approximately 7,000 gauss. The cross hatched material is lead. C_1, C_2, C_3, and C_4 are counters which control the simultaneous expansion of the two cloud chambers.

Figure 5. Source: J.C. Street, Journal of the Franklin Institute, 227 (1939), pp. 765-788.

density and curvature that indicated a mass equivalent to about 200 electron masses. The photograph subsequently has been reproduced in many texts.[133] (See Figure 6 below.)

This striking photograph was shown at many meetings. Furry recalls bringing a copy of it to England where in conferences it created quite an impression on the audiences; for many people it was the most impressive piece of evidence for the new particle. However, for Street and Stevenson the energy-range relation remained the conclusive evidence. In Street's words, this "was the approach I always considered the soundest and the most convincing but it wasn't the most convincing to the uninitiated listener. You had to have studied the subject and thought about it and figured out a way to have measured the mass. But it will convince you if you take the time to study it."[134] As for the photograph,

> The picture we did of a dense track near the end of its range was really just a thing you did for a demonstration lecture--it had to be there. It was just a question of getting it done . . . . [Still] my reason for not depending on it too much was that we never had but one or two of those [photographs] and anything can happen once, so we weren't too happy with our experiments on that.[135]

Convincing evidence for Street and Stevenson could only come by the large statistics garnered from the carefully calibrated counters set in coincidence circuits with the cloud chambers.

Figure 6. Reproduction of photograph by Street and Stevenson of stopping muon. Source: Rossi, Cosmic Rays, p. 105 from Street and Stevenson, Physical Review, 52 (1937), p. 1003.

6. Summary and Conclusion

The question posed at the outset of this essay, "When was the muon discovered?" has no unique chronological answer. Since almost all sea-level ionizing cosmic radiation is composed of muons, there is a sense in which the first eighteenth-century observer of a spontaneously discharging electroscope "discovered" the muon, since he was the first to directly observe their effect. With similar cogency, one could argue that Bothe and Kolhörster counter coincidence experiment indicated the passage of particles that in retrospect we know to be muons. And, of course, there is a sense in which Carlson and Oppenheimer discovered the muon, since they first suggested the existence of a particle of intermediate mass in 1936. Anderson and Neddermeyer first presented good data showing that energy loss measurements of shower particles fit the Bethe-Heitler theory. This implied (though we can see this only in retrospect) that the penetrating particles must be other than electrons. Or Street and Stevenson could be credited with the discovery for having shown that there was a characteristic difference between the shower producing power of shower particles and of penetrating particles.

Most authors now give the credit to Anderson and Neddermeyer's March 1937 energy loss argument or else to Street and Stevenson's April 1937 range-momentum argument. Both of these experiments definitively seemed to show the penetrating particles to fall into two distinct groups <u>in the same momentum range</u>. Of course, one equally well could attribute the discovery to Street and Stevenson's 1937 photograph of a stopping track, for this provided the first quantitative analysis of the actual particle mass.

There are many other points that could be chosen as the "moment of discovery." However, I hope to have shown that such a term--perhaps valuable for prize committees and physics textbooks--corresponds to little or nothing in the historical development of cosmic ray physics. The systematic investigation of the character of the rays began with the studies of Hess, and Millikan's interest in the origin and fate of the elements. As Pais remarked, when the muon was finally separated from the cosmic rays as a new kind of matter, the study of elementary particle physics had begun.

Instead of looking for a "moment of discovery," we should envision the evolution of cosmic ray physics as a progressively refined articulation of the properties of the rays. At each stage of the process, new characteristics could be attributed to them: they discharged electroscopes; the discharge rate varied in a certain fashion with depth in matter; the shower particles were more easily absorbed than the single particles; and so on.

Experiment and theory <u>both</u> contributed to this refinement of distinction. For example, without the cloud chamber photographs of Anderson, Street, Rossi, and others, Carlson and Oppenheimer would have had no basis on which to proceed with a theory of showers. Without the work of Carlson and Oppenheimer, the wedge between shower and penetrating particles could not have been as clear. And it was precisely the clarity of this distinction that made it manifest to Anderson, Street, and their co-workers that it was the penetrating particles that demanded an explanation, not the "electrons" of the showers accounted for by Bethe-Heitler theory. (See Figure 7.)

In the above I have stressed the closeness of experiment and theory for both the East and West Coast groups. Yet there were, despite their often parallel development, marked differences in their motivations, equipment, and styles of experimentation. On the West Coast, the systematic study of cosmic rays was originally fueled by Millikan's conviction that elements were being formed throughout space. This led, via $E = mc^2$ and $E = \nu h$, to his belief that photons at specific energy bands constituted the primary cosmic radiation. In turn, Millikan's theory of photons as primary cosmic rays determined his whole emphasis on the study of the absorption curves of the radiation.

Indeed, only by understanding Millikan's original project can one see the real origin of his experimental program of measuring discharge rates in lakes, on mountains, and under different thicknesses of lead. So too, can one only understand Millikan's mistakes such as his commitment to the "band theory," his persistent attacks on the reality of the latitude effect, and his assertion that there was an ionization maximum high in the atmosphere.

Along with Millikan and his co-workers' study of absorption curves came perhaps his greatest cosmic ray successes: his assignment to Carl Anderson of the problem of measuring the energy of the "secondary electrons." This too can only be understood as an offshoot of Millikan's original goals, for he fully expected Anderson to find corroborative evidence of bands. Finally, as Anderson progressed, discovering nuclear disintegrations and pair production, Millikan once again

WEST COAST

Millikan Birth Cry Theory (photons)

a) no latitude effect
b) ionization max. in atmosphere
c) no high-energy electrons
d) bands in electron energy spectrum
e) no penetrating particles
f) nuclear electrons, positrons

EUROPE, EAST COAST,...

Bethe, Rossi (corpuscles)

a) latitude effect
b) no ionization maximum
c) highly penetrating particles
d) East-West Effect
e) Dirac pair-creation

London, 1934: Rossi, Anderson, Neddermeyer, Bethe,....

Low energy shower particles as unexplained phenomena. High energy particles (electrons?) that do not obey Bethe-Heitler theory. <u>QED CONDEMNED</u>.

WEST COAST | ANDERSON ET AL.

June 1936 Anderson and Neddermeyer show specific energy loss approximates Bethe-Heitler for shower ptcles

EAST COAST | STREET ET AL.

March 1936. Street and Stevenson note shower particles produce many more showers than single particles

December 1936. Carlson and Oppenheimer show showers could be electrons obeying Bethe-Heitler theory. Penetrating particles would be "new to physics."

April 1937. Fussell conducts cloud-chamber tests with very thin plates checking Bethe-Heitler for showers.

March 1937. E/E energy loss measurements separate particles into two groups even at same momentum. Penetrating particles thus not electrons (Anderson and Neddermeyer)

April 1937. Energy-range relation shows penetrating particles of low energy that cannot be protons. (Street and Stevenson)

CONCLUSION: NEW PARTICLE OF INTERMEDIATE MASS EXISTS. <u>QED VINDICATED</u>

November 1937. Stopping track gives "muon" mass as approximately $130m_e$

Further developments include the resolution of the following questions: Are there many "muon" masses?, two mesons?, Yukawa particles? etc.

Figure 7. Summary of the discovery of the muon.

appropriated the results as yet further evidence for the Birth Cry of atoms.

By 1934, though, Anderson had begun to break away from Millikan's view and to entertain the emerging quantum theory of electrodynamics. Anderson's experimental techniques, however, remained altogether continuous with his earlier work. Millikan's assignment to Anderson in 1931 had been to measure electron energies. In demonstrating the existence of the positron, Anderson had made use of the energy difference of a particle above and below a lead plate. Finally, as he approached the problem of testing the predictions of quantum electrodynamics, he did so by improving his cloud chamber techniques for measuring energy losses of particles passing through lead. Clear photographs of tracks through lead plates indicating energy loss constituted persuasive evidence for Anderson. Both Anderson and Millikan remained suspicious of the statistical evidence afforded by counter experiments.

Quantum electrodynamics had come to cosmic ray physics in several ways. First, it had come through Bethe's work involving the calculation of cross sections for the interaction of photons and electrons for the passage of fast electrons through matter. Bethe had close contact with some experimentalists, especially with Rossi ever since Bethe's time in Italy. This was natural since Rossi's experiments were well suited to, indeed designed for, the investigation of corpuscular cosmic radiation. Similarly, in the United States, Furry (who was also working on quantum electrodynamics) consulted frequently with Street at Harvard. Furry's help, along with Street's longstanding interest in Rossi's work, contributed to the Harvard group's focus on the use of logic circuits and counter equipment to garner statistical evidence. Street found these statistical arguments much more compelling than the few individual photographs he and Stevenson later obtained. Years of experience with counter coincidence apparatus made the range-energy experiments persuasive to Street. Thus, while both the East and West Coast groups found their way to the muon, they did so with very different experimental styles. Consequently, they found different kinds of evidence convincing.

The two traditions finally converged in their 1937 conclusion that there existed a new particle, previously unknown, of mass intermediate between that of the electron and of the proton. More importantly, this convergence was at once a statement of theoretical and experimental physics. The discovery of the muon was inseparably bound up with the resurrection of quantum electrodynamics.

7. Epilogue

The acceptance of Anderson and Neddermeyer, and Street and Stevenson's new particle was greatly facilitated by an incorrect theoretical development. Yukawa had conjectured in 1935 that the nuclear forces might be due to the exchange of a heavy particle, by analogy to the exchange of the massless photon in quantum electrodynamics. Since the mu-meson had approximately the same mass as the Yukawa particle, Oppenheimer and others suggested they were in fact the same particle. A confirmation of this identification seemed to come in 1938-39 when observations of mu-meson flux at different altitudes indicated the mesons were decaying in flight approximately as they were predicted to do in Yukawa's theory. During the war, the decay of the mu-meson was confirmed in laboratory experiments.

Only after the war was over, when new methods of particle detection had been developed, were the mu-meson and the pion (the "true" Yukawa particle) distinguished by C. F. Powell. With this last development and the almost simultaneous transformation of quantum field theory by Feynman, Schwinger, and Tomanaga, elementary particle physics had begun.

After the discovery of the muon, Rossi, Street, and Anderson continued to work on cosmic ray problems. Millikan, however, never fully abandoned his Birth Cry theory. By 1939, he was forced to acknowledge the existence of cosmic ray particles too energetic to be accounted for by energy conversion from atom-building. At the same time, though, he inaugurated a new transformation in the cosmos to account for the cosmic ray particles, in which whole atoms were continually annihilated throughout the universe, producing high energy pairs of photons and electrons. Millikan's subsequent modifications of this last theory continued until his death, but all of them remained far outside the mainstream of modern physics.

NOTES

[1]G. Bernardini, talk given at the International Symposium on the History of Particle Physics, May 28-31, 1980, at Fermilab.

[2]J. A. Wheeler, "Some Men and Moments in Nuclear Physics," p. 242 in R. H. Stuewer, ed., Nuclear Physics in Retrospect (Minneapolis: Univ. of Minnesota Press, 1977).

[3]B. Rossi, Cosmic Rays (New York: McGraw-Hill, 1964), p. 109.

[4]J. C. Street, "Cloud Chamber Studies of Cosmic Ray Showers and Penetrating Particles," Journal of the Franklin Institute, 227 (1939), p. 778.

[5]Abraham Pais, "Particles," Physics Today, 21 (1968), p. 24.

[6]E. Rutherford and H. L. Cooke, "A Penetrating Radiation From the Earth's Surface," Physical Review, 16 (1903), p. 183; J. C. McLennen and E. F. Burton, "Some Experiments on the Electrical Conductivity of Air," Physical Review, 16 (1903), pp. 184-192.

[7]Cf. B. Rossi, Cosmic Rays, pp. 1-8.

[8]The Autobiography of Robert A. Millikan (New York: Prentice Hall, 1950), pp. 209ff.

[9]R. A. Millikan and I. S. Bowen, "High Frequency Rays of Cosmic Origin I. Sounding Balloon Observations at Extreme Altitudes," Physical Review, 27 (1926), pp. 353-361.

[10]R. A. Millikan and R. M. Otis, "High Frequency Rays of Cosmic Origin II. Mountain Peak and Airplane Observations," Physical Review, 27 (1926), pp. 645-658.

[11]Ibid., p. 658.

[12]R. A. Millikan and G. H. Cameron, "High Frequency Rays of Cosmic Origin III. Measurements in Snow-Fed Lakes at High Altitudes," Physical Review, 28 (1926), pp. 851-868.

[13]Ibid., pp. 860-861.

[14]J. H. Jeans, Problems of Cosmogony and Stellar Dynamics (Cambridge, 1919), p. 286; A. S. Eddington, The Internal Constitution of the Stars (Cambridge, 1926), ch. XI. Both cited in Millikan and Cameron, Proc. of the Nat. Acad. of Sciences, 14 (1928), p. 639.

[15]For more on early, pre-quantum mechanical ideas about fusion, see Joan Bromberg, "Concept of Particle Creation Before and After Quantum Mechanics," Historical Studies in the Physical Sciences, 7 (1976), pp. 161-191, esp. pp. 165 and 171ff.

[16]Millikan and Cameron, Physical Review, 28 (1926), p. 868.

[17]Ibid.

[18]Robert Kargon, "Birth Cries of the Elements: Theory and Experiment along Millikan's Route to Cosmic Rays," in The Analytic Spirit, H. Woolf, ed. (Ithaca: Cornell University Press, 1981).

[19]Robert W. Seidel, Physics Research in California: The Rise of a Leading Sector in American Physics. Unpublished Ph.D. dissertation,

University of California, Berkeley, 1978. See chapter VII, "Cosmic Rays."

[20]R. A. Millikan, Science and Life (Boston: The Pilgrim Press, 1924), p. 43. See also D. J. Kevles, "Robert A. Millikan," Scientific American, January 1979, pp. 142-151.

[21]The Autobiography of Robert A. Millikan (New York: Prentice Hall, 1950), p. 87.

[22]R. A. Millikan, Science and Life, p. 59. (Emphasis in original.)

[23]This kind of speculation about inorganic evolution was hardly unique to Millikan; Eddington, for instance, thought along similar lines. See, for instance, the discussion in J. Bromberg's note 39 in her "Concept of Particle Creation." Also see R. Kargon, "Birth Cries," reference in note 18 above.

[24]Ibid., pp. 27-28.

[25]R. A. Millikan, Evolution in Science and Religion (Port Washington: Kennikat Press, 1973), p. 81. Originally published in 1927 by Yale University Press.

[26]Ibid., p. 82.

[27]The petition is reproduced as an appendix to Millikan's Autobiography, pp. 289ff.

[28]R. A. Millikan and G. H. Cameron, "New Precision in Cosmic-Ray Measurements; Yielding Extension of Spectrum and Indications of Bands," Physical Review, 31 (1928), p. 922.

[29]G. Holton, The Scientific Imagination (Cambridge: Cambridge University Press, 1979), pp. 25-83.

[30]Ibid., pp. 78-79.

[31]Ibid.

[32]Millikan and Cameron, Physical Review, 31 (1928), p. 926.

[33]Ibid., p. 927.

[34]Ibid.

[35]R. A. Millikan and G. H. Cameron, "The Origin of the Cosmic Rays," Physical Review, 32 (1928), pp. 533-557.

[36]Ibid.

[37]R. A. Millikan and G. Harvey Cameron, "Evidence that the Cosmic Rays Originate in Interstellar Space," Proceedings of the National Academy of Sciences, 14 (1928), pp. 637-645.

[38]Ibid., p. 445.

[39]Millikan and Cameron, Physical Review, 32 (1928).

[40]Millikan and Cameron replaced the Compton formula (modified to give a mass absorption law for light) with an analogous mass absorption law based on Dirac's equation. Otherwise, the theory of absorption was the same.

[41]Millikan and Cameron, Physical Review, 32 (1928), p. 556.

[42]Letter from Oppenheimer to Millikan, 12 February 1929. Millikan microfilm, roll 23, file 22.1.

[43]Letter from Millikan to Oppenheimer, 11 March 1929. Millikan microfilm, roll 23, file 22.1.

[44]Interview of C. D. Anderson, 30 June 1966, by C. Weiner. Transcript at American Institute of Physics, p. 7.

[45]Reprinted in Nature, 127 (1931) and in the Annual Report of the Smithsonian Institution, 1931, pp. 277-285, as "Present Status of Theory and Experiment as to Atomic Disintegration and Atomic Synthesis."

[46]*Ibid.*, p. 278.

[47]*Ibid.*

[48]*Ibid.*

[49]R. A. Millikan and C. D. Anderson, "Cosmic-Ray Energies and their Bearing on the Photon and Neutron Hypotheses," Physical Review, 40 (1932), p. 327.

[50]*Ibid.*

[51]*Ibid.*

[52]Letter from Anderson to Millikan, 3 November [1931]. Millikan microfilm, roll 23, file 22.3.

[53]Anderson interview with C. Weiner, *op. cit.*

[54]C. D. Anderson, "The Apparent Existence of Easily Deflectable Positives," Science, 76 (1932), p. 238.

[55]C. D. Anderson, "Cosmic-Ray Positive and Negative Electrons," Physical Review, 44 (1933), pp. 406-416.

[56] R. A. Millikan, "New Techniques in the Cosmic-Ray Field and Some of the Results Obtained with Them," Physical Review, 43 (1933), pp. 661-666, on p. 662.

[57] Ibid. (Emphasis in original.)

[58] Anderson, Physical Review, 44 (1933), p. 415.

[59] Ibid., p. 411.

[60] R. A. Millikan, "Sur les Rayons Cosmiques," Annales de l'Institut Henri Poincaré, 3 (1932), pp. 447-464, on pp. 451-452.

[61] Ibid.

[62] New York Times, 5 February 1933, p. 1.

[63] A. H. Compton, "A Geographic Study of Cosmic Rays," Physical Review, 43 (1933), pp. 387-403.

[64] Letter from Millikan to Compton, 30 November 1932. Millikan microfilm, roll 23, file 22.18.

[65] Letter from Millikan to Victor Neher, 12 September 1936. Millikan microfilm, roll 24, file 22.15.

[66] W. Bothe and W. Kolhörster, "Das Wesen der Hohenstrahlung," Zeitschrift für Physik, 56 (1929), pp. 751-777, on p. 751.

[67] Author's taped interview with Bruno Rossi, 25 September 1980.

[68] B. Rossi, "Method of Registering Multiple Simultaneous Impulses of Several Geiger's Counters," Nature, 125 (1930), p. 636.

[69]Rossi's results are summarized in "Über die Eigenschaften der durchdringenden Korpuskularstrahlung im Meeresniveau," Zeitschrift für Physik, 82 (1933), pp. 151-178.

[70]Carl Anderson, R. A. Millikan, Seth Neddermeyer, and William Pickering, "The Mechanism of Cosmic-Ray Counter Action," Physical Review, 45 (1934), pp. 352-363, on p. 352. (Emphasis in original.)

[71]Ibid.

[72]P. M. S. Blackett and G. P. S. Occhialini, "Some Photographs of the Tracks of Penetrating Radiation," Proceedings of the Royal Society of London, 139A (1933), pp. 699-720.

[73]See also Millikan, Physical Review, 43 (1933), p. 663.

[74]See also articles by Wheeler and Bethe in Nuclear Physics in Retrospect (reference 2), and L. Brown, "The Prediction and the Discoveries of 'Yukawa's Meson'. Part 1. The Prediction," Centaurus, (1981) pp. 71-132.

[75]J. Bernstein, Hans Bethe, Prophet of Energy (New York: Basic Books, 1980), p. 20.

[76]Ibid., p. 25.

[77]Annalen der Physik, 5 (1930), pp. 325-400.

[78]Ibid., pp. 325-326.

[79a]Bohr's semi-classical treatment of the energy loss of an electron passing through matter was developed in, N. Bohr, "On the Theory of the Decrease of Velocity of Moving Electrified Particles on passing

through Matter," *Philosophical Magazine*, *25* (1913), pp. 10-31, and in N. Bohr, "On the Decrease of Velocity of Swiftly Moving Electrified Particles passing through Matter," *Philosophical Magazine*, *30* (1915), pp. 581-612. For a discussion of the role of collision theory in the construction of the Bohr atom, see John Heilbron and Thomas S. Kuhn, "The Genesis of the Bohr Atom," *Historical Studies in the Physical Sciences*, *1* (1969), pp. 211-290.

The idea of Bohr's work had been to treat the passing particle's electric field by a radiative field of the same Fourier components. Each of the Fourier components could then be separately examined in its effects on the stationary atom's electrons. This method was extended to a relativistic treatment by Bohr, E. J. Williams, and C. F. v. Weiszäcker, making use of some results on photon-photon interactions by Breit and Wheeler. This treatment of charged particles passing through matter using the impact parameter (and not, as Bethe did, the momentum transfer) has become known as the Weiszäcker-Williams method. The fact that the authors were successfully able to extend the Bohr approach to the relativistic domain gave added confidence to theorists that Bethe's approach was correct. See the excellent review of this line of work in J. A. Wheeler's article cited in reference 2, esp. pp. 242ff. Also for a modern discussion of the relation of the various stopping formulae and approximations, see S.P. Ahlen, "Theoretical and Experimental Aspects of the Energy Loss of Relativistic Heavily Ionizing Particles," *Reviews of Modern Physics*, *52* (1980), pp. 121-173.

[79b]Author's taped interview with H. Bethe, 11 December 1980.

[80]H. Bethe, "Bremsformel für Elektronen relativistischer Geschwindigkeit," Zeitschrift für Physik, 76 (1932), pp. 293-299.

[81]Bernstein, p. 38.

[82]Bethe Interview.

[83]Heitler's work first presented in print in Zeitschrift für Physik, 84 (1933), pp. 145-167; and then with F. Sauter in Nature, 132 (1933), p. 892.

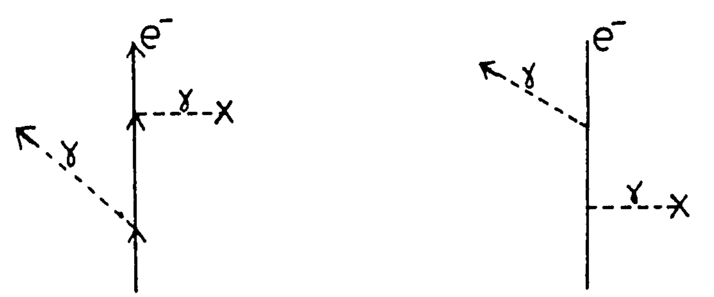
[84]W. Heitler and F. Sauter, "Stopping of Fast Particles with Emission of Radiation and the Birth of Positive Electrons," Nature, 132 (1933), p. 892.

[85]Letter: Heitler to Bohr, 16 October 1933, BSC roll 20. BSC = Bohr Scientific Correspondence, microfilm copy on deposit in the American Institute of Physics.

[86]Bethe Interview.

[87]"On the Stopping of Fast Particles and on the Creation of Positive Electrons," Proceedings of the Royal Society, A 146 (1934), pp. 83-112.

[88]In modern terms, the two terms correspond to the two Feynman diagrams,



[89]Bethe and Heitler, Proceedings, p. 103.

[90]*Ibid.*, pp. 104-105.

[91]C. D. Anderson and S. Neddermeyer, "Energy-Loss and the Production of Secondaries by Cosmic-Ray Electrons," Physical Review, 46 (1934), p. 325.

[92]"Fundamental Processes in the Absorption of Cosmic-Ray Electrons and Photons," International Conference on Physics London 1934 (Cambridge: Cambridge University Press, 1935), p. 181. Privately, however, Anderson offered Bethe somewhat more conciliatory judgment of the agreement between theory and experiment. In a letter to Bethe of 7 June 1935, Anderson remarked that his experiments with Neddermeyer, "so far are very incomplete and not accurate, and there is not a great deal that can be said about the validity of the theoretical formulae. It looks as though the theory and experiment do not conflict badly for electron energies below 100 MeV, but that for higher energies the formulae give too high values for the absorption."

[93]Anderson and Neddermeyer, London, p. 181.

[94]*Ibid.*, p. 182.

[95]At low energies, the results of Chadwick, Blackett and Occhialini and the early results of Anderson were now seen by Anderson as "combin[ing] to show the success of the Dirac theory as developed by Oppenheimer and Plessett and by Heitler and Sauter in interpreting the results obtained for photon energies of 2.6 MeV." London, p. 183. Note the contrast with Anderson's earlier papers where he criticized the Blackett and

Ochialini work and supported the Millikan "Birth Cry Theory" of nuclear electron emission.

96 _London_, p. 250.

97 Letter: Weiszäcker to Bethe, 5 December 1934, Bethe Archives. Bethe's "reversal" was this: soon after the conference, he and Compton concluded the latitude effect, east west effect, and penetrating particles all could be accounted for if one supposed the charged cosmic rays to predominantly include predominantnly protons. See A. H. Compton and H. A. Bethe, "Composition of Cosmic Rays," _Nature_, _134_ (1934), pp. 734-735.

98 A. Smith and C. Weiner, _Robert Oppenheimer, Letters and Reflections_ (Cambridge: Harvard Univ. Press, 1980), p. 161.

99 Letter: Oppenheimer to Bohr, 14 June 1933, in _Oppenheimer_, pp. 161-162.

100 _Ibid_., p. 164.

101 Letter: Oppenheimer to Uhlenbeck, Fall 1933. _Ibid_., pp. 167-168.

102 _Ibid_., p. 181.

103 J. R. Oppenheimer, _Physical Review_, _47_ (1934), pp. 44-52.

104 _Ibid_., p. 45.

105 "On the Quantum Theory of the Electromagnetic Field," _Proceedings of the Royal Society_, _143_ (1934), pp. 410-437.

[106]W. H. Furry and J. F. Carlson, "Production of High Energy Electron Pairs," Physical Review, 45 (1934), p. 137.

[107]Letter: Bohr to Williams, 11 February 1935, BSC Roll.

[108]Anderson, Early Work, p. 828.

[109]Author's taped interview with J. C. Street, October 1979; Rossi Interview.

[110]Street Interview.

[111]Rossi, Cosmic Rays, pp. 69ff.

[112]Street Interview.

[113]Ibid.

[114]Ibid.

[115]E. C. Stevenson and J. C. Street, "Nature of the Penetrating Cosmic Radiation at Sea Level," Physical Review, 47 (1935), p. 643.

[116]Ibid., p. 800.

[117]J. C. Street and E. C. Stevenson, "The Absorption of Cosmic-Ray Electrons," Physical Review, 47 (1935), pp. 891-895.

[118]J. C. Street and E. C. Stevenson, "Design and Operation of the Counter Controlled Cloud Chamber," Physical Review, 49 (1936), pp. 425-428.

[119]Ibid.

[120]C. D. Anderson and Seth H. Neddermeyer, "Cloud Chamber Observations of Cosmic-Rays at 4300 Meters Elevation and Near Sea-Level," Physical Review, 50 (1936), p. 263.

[121]J. F. Carlson and J. R. Oppenheimer, "On Multiplicative Showers," Physical Review, 51 (1936), pp. 220-221. Shower calculations were independently done by H. J. Bhaba and W. Heitler, "Passage of Fast Electrons through Matter," Nature, 138 (1936), p. 401.

[122]Diagrams are purely illustrative and are not to be taken as Feynman diagrams.

[123]Carlson and Oppenheimer, Physical Review, p. 220.

[124]Ibid., p. 221.

[125]S. H. Neddermeyer and C. D. Anderson, "Note on the Nature of Cosmic-Ray Particles," Physical Review, 51 (1937), p. 884.

[126]Ibid., p. 886.

[127]The authors explain the "negative energy loss" as being attributable firstly to statistical spread and secondly to upwardly travelling particles.

[128]L. Fussell, "Production and Absorption of Cosmic-Ray Showers," Physical Review, 51 (1937), pp. 1005-1006.

[129]Ibid.

[130]J. C. Street and E. C. Stevenson, "Penetrating Corpuscular Component of the Cosmic Radiation," Physical Review, 51 (1937), p. 1005.

[131]Street Interview.

[132]Bethe Interview; Furry Interview.

[133]For example, see Rossi, Cosmic Rays, p. 105.

[134]Street Interview.

[135]*Ibid*.

[136]See L. Brown, "Yukawa's Meson," cited in note 74.

CHAPTER IV

THE DISCOVERY OF WEAK NEUTRAL CURRENTS AND THE TRANSFORMATION OF EVIDENCE

1. Introduction

Over the course of a year and a half, from the fall of 1972 to the spring of 1974, photographs like those reproduced in Figures 1 and 2 were transformed from curiosities to the principal evidence for the existence of weak neutral currents. As such they came to be seen as the first concrete experimental support for the gauge theories of Weinberg and Salam. Since that time, gauge theories in a more general form have lent hope to the possibility of extending the unification to the strong interaction as well. It is the goal of this paper to approach the experimental discovery of weak neutral currents from a historical perspective.

By historical, I do not mean a review of the kind that usually appears in the Reviews of Modern Physics and other journals which have as their principal aim the logical and coherent reconstruction of a sequence of developments in physics. This has been done for neutral currents by, among others, A. K. Mann, A. Rousset, C. H. Llewellyn Smith, C. Baltay, G. Myatt, J. E. Kim et al., and D. Cline and W. F. Fry.[1] Rather, I have in mind the kind of studies historians have applied to events that occurred further in the past.

The history of the discovery of neutral currents must be traced in parallel, on both sides of the Atlantic. Work in the United States was centered at the National Accelerator Laboratory (often referred to as NAL or Fermilab) in Batavia, Illinois, by a collaboration of groups

Figure 1. Neutral current event from CERN-Gargamelle. All tracks from the vertex stop or decay in liquid indicating they are hadrons and not muons.

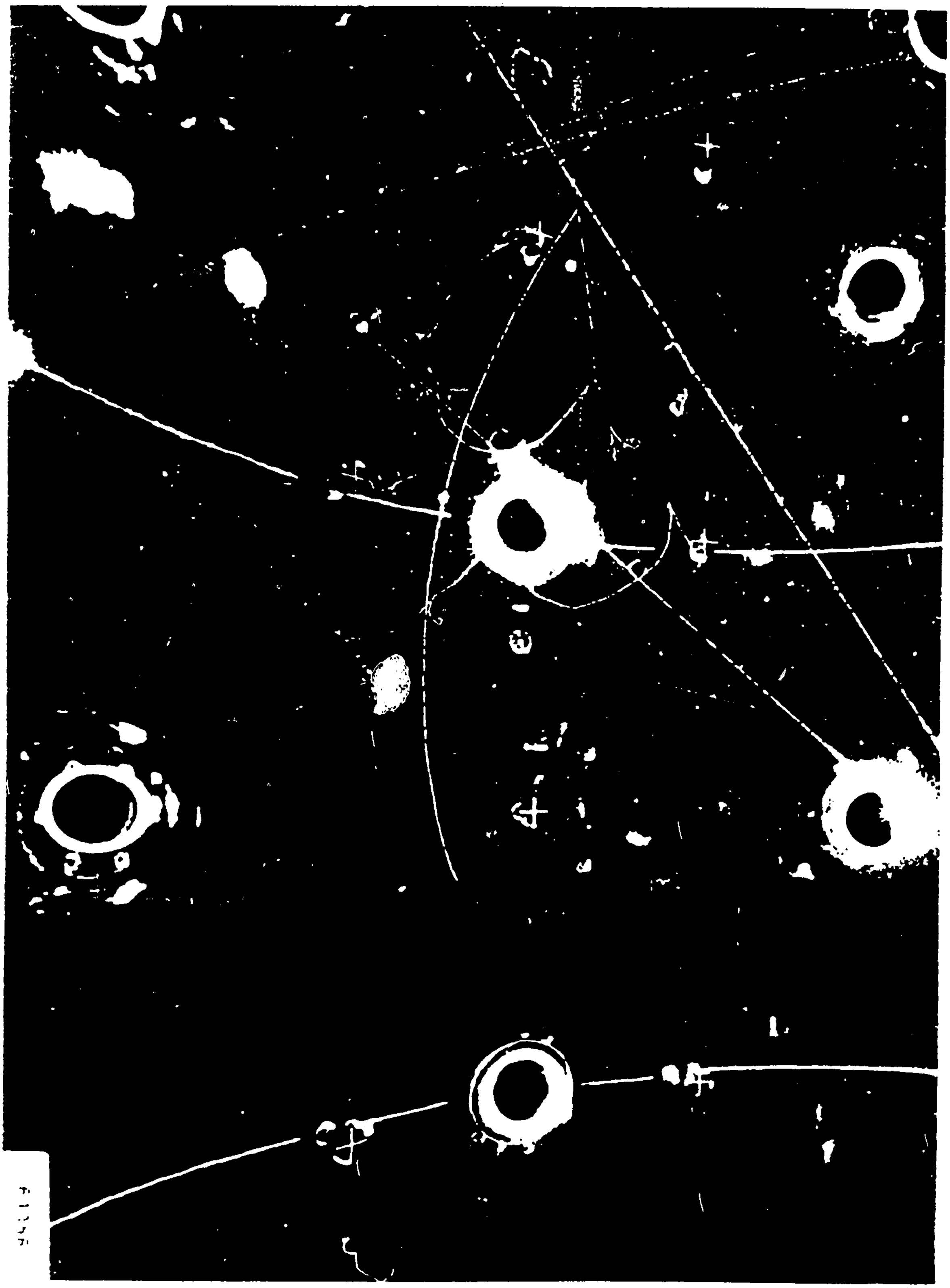
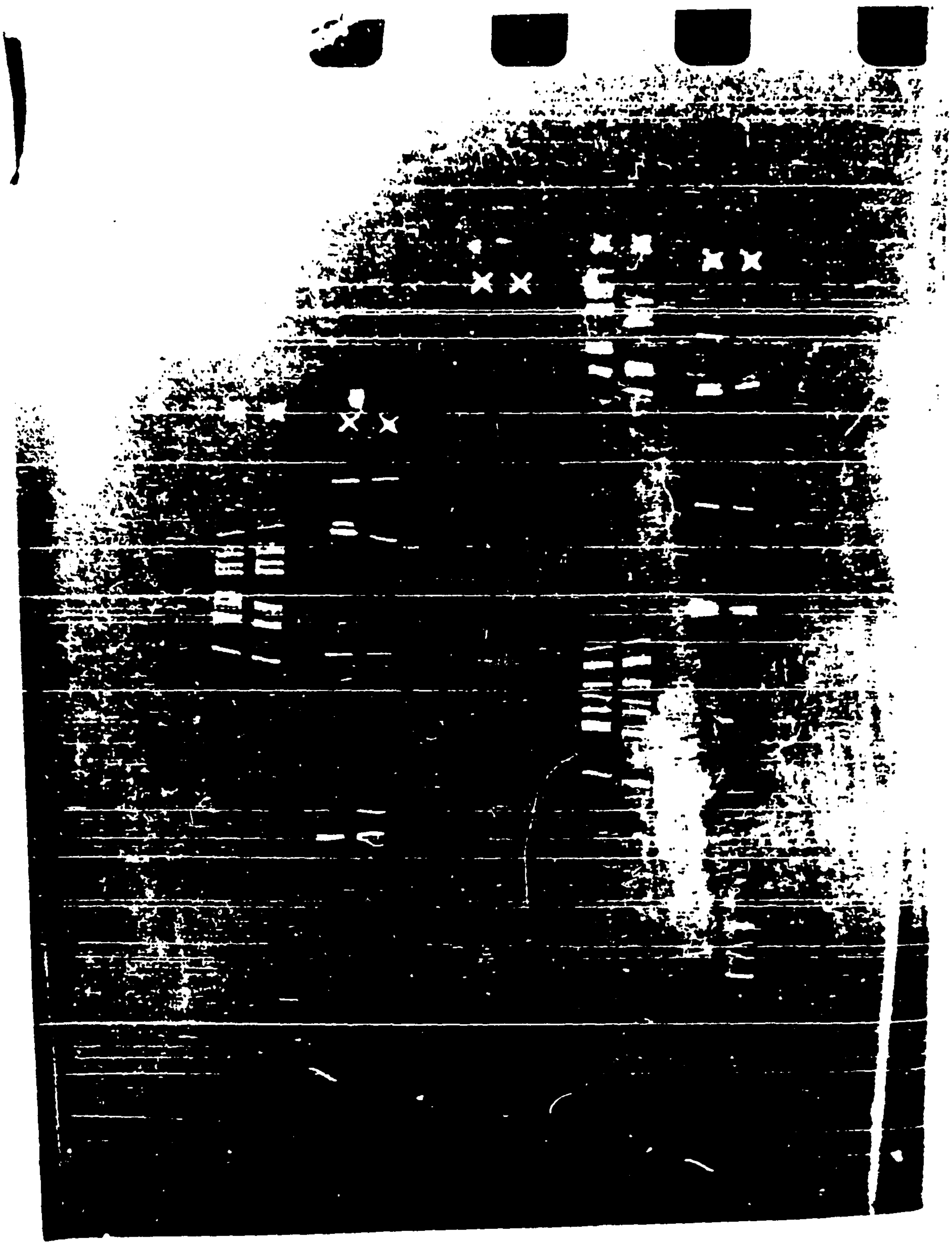


Figure 2. Neutral current event from NAL experiment 1A. Two vertices are stereo view of one event. Again, hadrons are present with no muons.



representing Harvard, Wisconsin, Pennsylvania, and Fermilab. In Europe, the search for neutral currents was performed primarily at CERN (Centre Européen de Recherche Nucléaire) outside Geneva by laboratories from France, England, Italy, Germany, Belgium, and CERN itself. Over seventy-five physicists eventually signed the various early reports on neutral currents; many other experimentalists and theorists participated in innumerable stages of planning, discussion, analysis, and interpretation that began long before the experiment itself and continued long afterwards.

By the time the experiments neared completion, it had become clear to a great number of physicists, participants, and non-participants alike, that much was at stake. Even though gauge theories had also been constructed without neutral currents (notably by S. L. Glashow and H. Georgi), it is clear that the NAL and CERN confirmation of the Weinberg-Salam prediction was instrumental in creating a climate of acceptance for the specific model of Weinberg and Salam, as well as, more generally, for the broad approach of gauge theories with spontaneously broken symmetries.

The experimental and theoretical consequences of these neutral current experiments provide one rationale for examining them historically; but beyond the larger framework of the history of gauge theories, the events surrounding the discovery of neutral currents present a particularly interesting opportunity to study the relation of experiment to theory in elementary particle physics. In addition, the history of neutral currents also raises a number of methodological issues for the historian: Can the tools of history of physics, as they have been developed for the study of earlier periods, such as the scientific revolution, statistical mechanics, relativity, and quantum mechanics, be applied to events which are or are almost contemporary with us? Must events be ten or twenty years distant before a historical perspective is possible? Finally, are the tools developed for the study of the history of quantum mechanics (for example) appropriate to a history of events half a century more recent?

This analysis presupposes that a contemporary history of physics is possible, but with an important caveat: the archival record has a different character from the record of even thirty years ago. There is no more detailed scientific correspondence of the sort that forms the backbone of the classic works of the history of physics from Descartes to Einstein and Bohr. Still, some of the older types of sources remain such as draft manuscripts, transcripts of conference meetings, and occasionally notebooks. There also are a variety of new sources which previously have not existed: preprints, computer simulations, computer calculations, internal memoranda, minutes of collaboration meetings, log books, and detailed experimental and grant proposals. I have used such sources in addition to the more usual ones, along with about twenty-five hours of interviews conducted in the United States, France, England, and Switzerland.

With these tools, I hope to offer an account of how two high-energy physics experiments--admittedly exceptional ones--took place. In some respects, as we will see, these experiments structurally resemble much older experiments as they have been documented. In other respects, amid the new technology, organization, and subject matter of high-energy physics, something quite new has entered the picture in recent decades.

2. The CERN Experiment "Gargamelle": From W-Search to Neutral Current Test

It is beyond the scope of this study to review the development of weak interaction theory up to the early 1960s. However, from the experimentalists' point of view, the theoretical interests were woven together concisely into a broad experimental program outlined by M. Schwartz, and by T. D. Lee and C. N. Yang in two consecutive (and related) Physics Review Letters.[3] A number of their suggestions became the guiding principles for experiments that were conducted over the next decade, including the Schwartz, Steinberger, and Lederman two-neutrino experiments,[4] tests of the Conserved Vector Current Hypothesis, lepton conservation, and electron-muon universality. Most important for our interest in neutral currents was their suggestion that neutrino interactions could be used in the search for the intermediate vector boson, W. At the time there were two main reasons to believe the W might exist: First, it would base weak interactions on an exchange force similar to the successful quantum electrodynamics, and second, it offered at least a hope that the theory could then be renormalized.[5]

In their own work on the W published that same year,[6] Lee and Yang concluded that the search for charged W's could be made by studying the reaction

$$(1) \qquad \nu + Z \longrightarrow W^{+} + \ell^{-} + Z.$$

Furthermore, since kaons did not undergo a fast decay to $W + \gamma$, they concluded that the mass of the W must be greater than that of the kaon;

from the form of the V-A interaction they could say the $W^{\pm}$ must be of spin 1. Beyond that little was known. The search suggested by Yang and Lee assumed that the W has approximately the mass of the heaviest known particles, nucleons. With this assumption, they could make a rough calculation of the cross section of reaction (1), and the rate of production of the W decay products, $\mu + \nu_{\mu}$. "If experimentally no $W^{\pm}$ is found," they wrote, "it would be possible to set a lower limit on the value of m_W."[7]

The search for $W^{\pm}$ formed one of the main motivations for the construction of Experiment 1A at the National Accelerator Laboratory and of Gargamelle at CERN. By contrast, the search for neutral currents had low priority since there seemed to be no pressing reason for neutral currents to exist in the phenomenological theory. Thus, when in February of 1964, A. Lagarrigue, A. Rousset and P. Musset assembled a preliminary project proposal for a new bubble chamber, their interest was centered on the search for <u>charged</u> intermediate vector bosons,[8] even though the phenomenological theory put no constraints on the mass of the W. Consequently, there was no assurance that the neutrino energies in the new proposal would be sufficient to produce the W. Still, as the Sienna conference of 1963 approached, it was hoped that the particle was somewhere in the range of few GeV, i.e., within the grasp of the next generation of experiments.[9]

Other projects were mentioned by Musset and Rousset, but even as the draft went through several revisions, neutral currents were not mentioned. Some six years later, with Gargamelle nearing completion,

D. H. Perkins again rewrote the proposed physics program for Gargamelle.[10] At the top of the list of physics projects remained the search for the intermediate vector boson. Then came the study of various processes predicted from the older theory of weak interactions which did not use W's. Still, not everything was the same. Since the original project proposals had been submitted, a group at SLAC had conducted an experiment in which they inelastically scattered electrons on nuclei with a large transfer of momentum.[11] The astonishing result of the SLAC work was that the double differential cross section ($d^2\sigma/dE'd\Omega$), when compared with the simple point Mott scattering, was independent of momentum transfer and center of mass hadron energy values up to 5 GeV.

The constant value of the ratio $d^2\sigma/dEd\Omega)/(d\sigma/d\Omega)_{Mott}$ suggested to Feynman and Bjorken that the electrons in fact were scattering from point scatterers that, in a not yet clear sense, were included in the protons and neutrons. Feynman christened these point scatterers partons. (These were first identified with quarks by Bjorken and Paschos.)[12] Tests of other consequences of the parton model suddenly became of exceptional interest.

As a result of this new interest, there was an added incentive to study the behavior of the neutrino interactions at high energies. If the SLAC results held good at even higher energies, the neutrino cross section was expected to rise linearly with energy. While the search for the W was now accompanied by excitement over the parton model, neutral currents remained relegated to a secondary place. "In addition," Perkins wrote, "there are of course, many other topics of interest for example, neutral currents, ... However, these problems can also be investigated with [other] chambers. On the other hand Gargamelle is, we claim, a unique instrument for investigating problems [like the W and the parton hypothesis]."[13]

During 1971, a great number of tests had to be completed in Gargamelle, and a physics program began to take shape. The analysis of these early photographs revealed some unusual events. As J. P. Vialle later recalled:

> One thing we saw right away on the photographs was that there were very energetic events in the bubble chamber with no muons. But obviously, we couldn't have said these are neutral currents, no, it wasn't like that. I think our first thought was that it was curious we were observing stars that could come from neutrons but were much too energetic, and that we would have to look into that.[14]

The experimental program continued as planned, with the study of neutral currents of secondary importance. Suddenly, in the spring of 1971, theoreticians began to take a new interest in neutral currents.

Ever since 1967, when it was first put forward, the Weinberg-Salam theory[15] had played absolutely no role in the experimentalists' planning. One reason for this was that t'Hooft's proof that the theory was renormalizable came almost four years after Weinberg's original paper.[16] Once the proof was made known and accepted, theorists such as E. A. Paschos, L. Wolfenstein, A. Pais, S. B. Treiman, Weinberg, and t'Hooft began to calculate some of the experimental consequences of the theory.[17] A combination of the renewed theoretical interest and the availability of cross sections they could test awakened the interest of the experimental community.[18]

Not long after the appearance of t'Hooft's paper in November of 1971, B. Zumino spoke to a group of experimentalists and theorists in the small library room in the building that housed Gargamelle at CERN. He explained to them the sudden interest in the now renormalizable Weinberg-Salam theory. Musset recalls being a little discouraged at the test the theorists were most in favor of: scattering a muon neutrino off an electron. Though extremely "clean" of background effects, the cross section (or likelihood) of such an event was extremely small. By contrast, the hadronic weak neutral current reaction seemed to have a chance of having a much larger cross section by analogy with the charged current cross section. Unfortunately, the theorists felt any calculation involving hadrons and the strong interaction would become much too complicated.

Musset's enthusiasm for the hadronic channel led him to conduct a literature search of earlier work on the subject.[19] In retrospect, it is perhaps fortunate that he did not find certain works which might have discouraged him completely. For instance, he did not know about the early CERN work of 1963 in which some of his own senior colleagues (Perkins and D. C. Cundy) had participated: they had published upper limits on the process Musset was considering which were far below the Weinberg level, below 3%.[20]

Musset recalled that his suggestion that the group study the hadronic neutral currents was met with no great enthusiasm by some other members of the collaboration.[21] Their lack of enthusiasm was certainly not because of any lack of interest in the question of neutral currents. Cundy, Perkins, H. Wachsmuth, and G. Myatt, for instance, had participated in the 1963 measurement; in 1970 they all participated in setting another less stringent bound on the ratio of hadronic neutral current to charged current events (NC/CC) = $.08 \pm .04$.[22] Indeed, it was precisely by their earlier experience that many of the Gargamelle collaborators knew at first hand the extreme difficulty of extracting any information on neutral currents from the background. The problem was that the neutrinos from the beam inevitably caused a large but unknown number of neutrons to enter the chamber from the surrounding magnets, floor and structure. If one of these secondary neutrons then hit a neutron or proton in the bubble chamber, the resulting shower of hadrons could look like a genuine neutrino neutral current event. In both cases no muon would emerge.

Gargamelle was much bigger than any previous bubble chamber; for this reason, it alone provided the opportunity to determine the rate of neutron background. For it was known at the time that neutrons had an interaction length in the bubble chamber liquid longer than the dimensions of the older bubble chambers. This meant that there was no way in the old experiments to see the exponential decrease of neutron-induced events as one looked further from the walls; hence it was impossible to figure out precisely how problematic neutron events would be. In Gargamelle, by contrast, not only could one see the exponential decay, one could do better by actually studying the entire career of a neutron in the liquid by examining the so-called "associated events."

Associated events look like this:

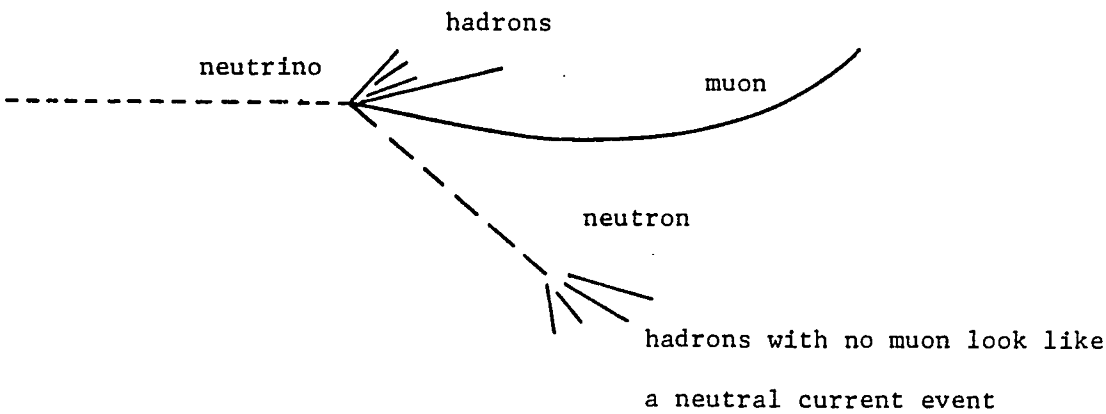


Figure 3. Associated Event

Upstream they have a normal charged current event from which a neutron is emitted, creating within the visible volume of the bubble chamber a "fake" neutral current event. By studying the length and angle of the neutron's path, one could then program the computer simulation to describe such events even where one did not see the neutron's beginnings. Unfortunately, especially at the beginning of the experiment, the associated events were quite rare.

In addition to the study of associated events, Musset, Pullia, and Lagarrigue, insofar as possible, attempted to treat the charged and and "neutral" events on equal footing. That is, criteria for selection of the hadronic part of the neutron interactions (location, energy, etc.)

were chosen to be precisely the same for charged and muonless events. Furthermore, since the primary question at the time was whether neutral currents existed (not yet in what proportion), only completely unambiguous events were used so as not to confuse charged with neutral currents. Finally, to reduce the effects of any remaining biases, and to make the measurement less sensitive to flux calculations, the group chose to focus attention on the ratio of neutral to charged currents.

These innovations proved crucial in demonstrating the existence of neutral current events, because in muonless events a large fraction of energy is carried off by the (unseen) neutrino. (About the same fraction is carried off in charged current events by the muon.) Thus, if one naively compared the total visible energy deposited for both charged and muonless events, one would then be measuring charged current events' energy by hadron energy plus muon energy, and neutral events by hadron energy alone. Therefore, since the number of events of both types falls off very quickly with energy, at a given total visible energy one would find an extremely and artificially low ratio of neutral to charged events. The failure to treat both kinds of events by hadron energy alone may well have been partially responsible for some earlier experiments' mistakenly low upper bound on the rate of neutral current events.

3. Background and Signal

By April 1972, Lagarrigue considered the search for neutral currents to be one of the three primary goals of the neutrino program for, as he wrote to Jentschke, then Director General of CERN, "following Weinberg's theoretical publication, everyone is anxious to discover whether neutral currents really exist."[23] The search for neutral currents,which had begun in January 1972 at Gargamelle, had thus become by late spring of 1972 one of the important areas of investigation for a number of physicists in the CERN group. Individuals disagreed, however, as to whether they thought the experiment would confirm or refute Weinberg's theory, and as to whether they thought the leptonic or hadronic channels should be pursued preferentially. By the spring, for example, Baltay and Cundy concentrated almost exclusively on the single electron search.[24]

During this time Perkins was in Oxford, where he composed a technical memorandum that was sent to the Gargamelle collaboration. Its object was, as it stated,

> to encourage people in the collaboration to study carefully the question of neutral hadronic currents, . . . because (i) there probably is an appreciable effect, (ii) conditions for proving the existence of neutral currents are much more favorable [than in the old bubble chambers at CERN] . . . . This is a big effect, large enough that a detailed and systematic analysis in Gargamelle, using the position of inter-

> actions in the chamber as well as the much better statistics of events, should be able to demonstrate, for the first time, the existence of neutral currents.[25]

Charles Baltay wrote back criticizing the memorandum, arguing that excess NC events could be accounted for by low energy muons alone, a position Perkins took issue with.[26] Responding to Baltay's criticism, Perkins argued that the data for the old bubble chamber experiment showed very few low energy muon events. "To summarize," Perkins wrote, "I don't think your explanation works, and I still cannot account for the excess neutral events, although I am certainly not going to claim that they prove the existence of neutral currents. . . . In providing the final solution (if there is one), one certainly needs to find a satisfactory explanation of the old data."[27] As did several other members of the collaboration, Baltay remained an enthusiast of the leptonic search which was cleaner. He continued this work later at the 15' FNAL chamber.

Musset, Pullia, and others at Milan continued work on the hadronic channel, and in June, Pullia presented a progress report on the hadronic neutral currents, suggesting that the neutron problem could be solved but offering no definite opinion on whether a significant level of neutral currents would remain when the background was removed.[28]

Sufficient interest in the neutral current question had developed by this time for the group to meet by itself in Paris, apart from the rest of the neutrino collaboration, in order to prepare a report for a conference at Batavia in September.[29] Before the meeting in Paris, the organizers requested that photographs of all neutral current "candidates" be sent ahead to CERN. By doing so, the authors of the memorandum hoped to standardize the criteria used to separate charged from neutral events. The data cards would be processed in order to plot energy of the events against the total longitudinal and transverse momenta, as well as against position. From these data the group hoped to get a first glimpse of the background problem with some statistical significance. Unlike Pullia's report, the hadron group now dropped the search for the relatively rare pion-producing events and decided to concentrate instead on determining the total (inclusive) cross section for $\nu_\mu + N \longrightarrow \nu_\mu +$ (hadrons), which was much larger.

Along with the proposal for a meeting, the first memorandum dealing purely with the subject of neutral current was issued.[30] At the very beginning of this report, the group noted that their best chance of isolating the neutrino from neutron interactions would be at high energies (the neutron spectrum was peaked below that of the neutrinos). By this time, then, the hadronic group's effort was entirely concentrated on the background problem which they described as five-fold:

(1) particles entering the chamber with the beam and which interact in the chamber;

(2) neutrons or kaons coming from outside chamber generated by neutrinos;

(3) cosmic rays;

(4) μ^-'s sufficiently slow to stop in the liquid;

(5) K_2^0's whose interaction length might be greatly extended by regeneration effects.[31]

Not everyone in the group was equally worried about all of these problems. For instance, Fry was especially concerned about the possibility of K_2^0 regeneration, no one was especially worried about cosmic rays, and everyone was interested in the slow muon and neutron problem.

The authors argued that the problem of stopping muons (which look like hadrons) could be attacked in several ways. Their number could be estimated from the scaling hypothesis, but this was considered a bit tenuous since the center of mass hadron energies were much higher than those studied at SLAC. All short unidentifiable tracks could be discarded, or finally, an upper limit to the number of "hidden μ contamination" could be set as follows. The muon spectrum had been measured in the liquid as had been the decay rate. From these facts, the number of muons below a certain energy in Gargamelle could be calculated. Then from the theoretical ratio (of muon capture/muon decays) the number of muons captured below a certain energy could be found.

Lastly, the group set up a standardized system for the recording of the neutral current candidates. At least one physicist would review each event and bit by bit the data would be assembled in preparation for Batavia. At Batavia, Perkins summarized the group's work from <u>his</u> perspective. It is worth quoting from his assessment of the future prospects of the hadronic and electron neutrino experiments:

> As far as the Weinberg theory is concerned, the most definitive and unambiguous evidence for or against, must come from the purely leptonic reactions . . . since the hadronic processes involve details of strong

> interactions which might contain unknown suppression effects . . . . As I have tried to indicate, the reactor experiment [$\nu_e + e \rightarrow \nu_e + e$] is beset with severe background problems. Even if in [other] future improved experiments, a clean signal is detected, it is necessary, in order to finally demolish the Weinberg theory to prove that the observed signal rate is consistent with the V-A predictions within close limits. It is difficult to believe that this could be achieved to a precision of better than 20%.[32]

By contrast, Perkins pointed out, certain purely leptonic interactions could occur only in a theory other than the old phenomenological V-A theory. This was a process Perkins felt the Gargamelle group could proftably investigate: "In the CERN Gargamelle experiment to date, the expected number of events was between 1 and 9, and none was observed. . . . If none were observed [in the remainder of the experiment], this would be fairly conclusive evidence against the Weinberg theory."[33] It would seem that the division in the CERN collaboration, which many of the members recalled very vividly,[34] was based not on whether neutral currents should be searched for, but in which process.

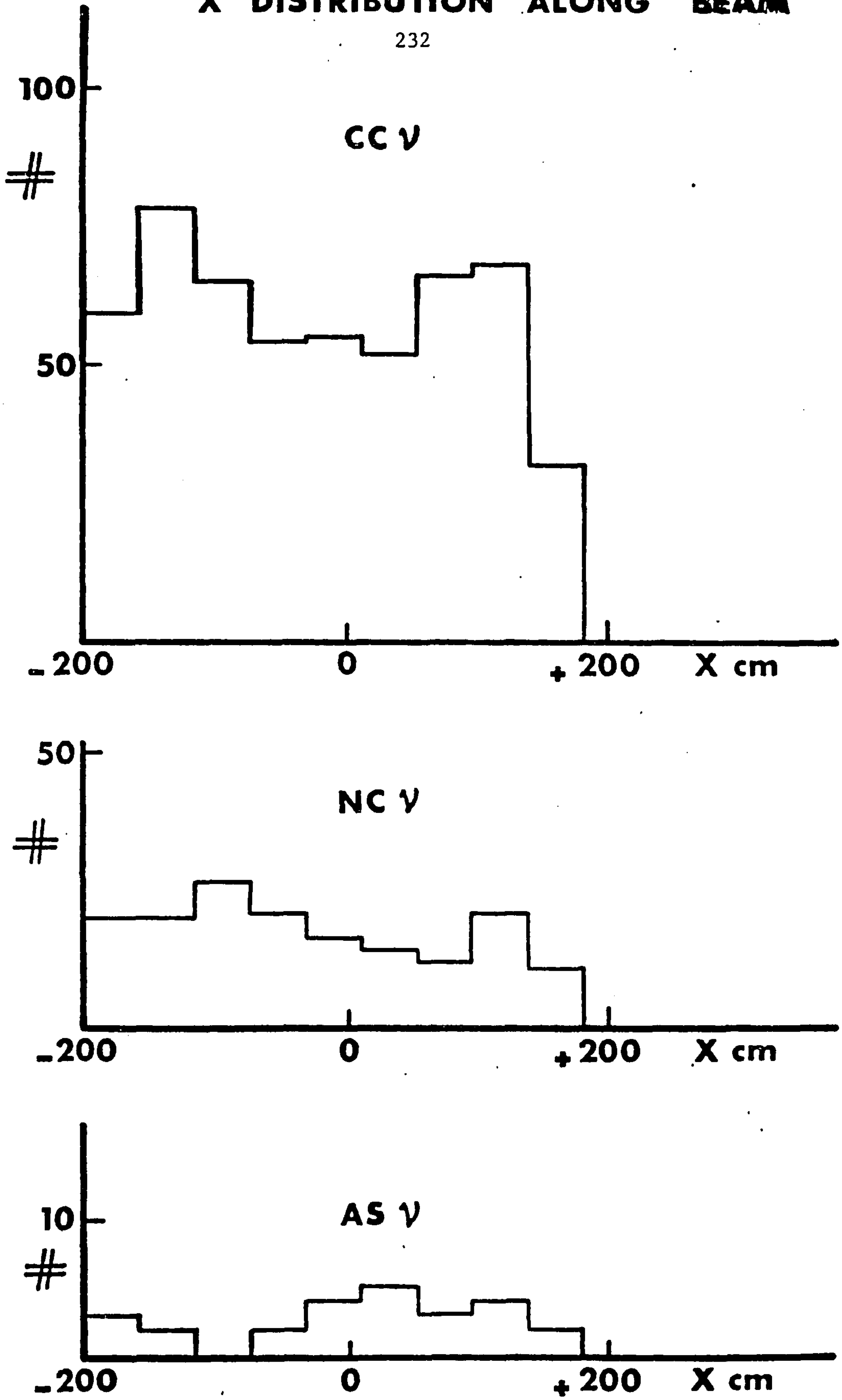
The fall of 1972 was spent with some groups conducting their respective data analyses, a long, often frustrating task in which hundreds of events had to be compared, definitions modified, and criteria adjusted. By January 1973, Musset and the others had gathered sufficient data to present their findings to the American Physical Society meeting in New York.

The emphasis of Musset's talk was almost entirely on the neutron background problem, and the data presented was in the form of number of events (charged, neutral, associated) plotted against longitudinal and radial position in the bubble chamber. His goal was to demonstrate that the events occurred relatively evenly throughout the volume of the machine, as would real neutrino events as opposed to neutron-induced events. This was not the most reliable check, but there were at the time very few associated events to study, because as the group had pushed up the minimum required energy of the hadrons in order to exclude neutron-induced events (which tended to have lower energy) they eliminated the vast majority of their data. The cut-off especially reduced the number of associated events. They therefore had to rely entirely on the spatial distribution of events. (See Figure 4.)[35]

The data were beginning by this time to indicate that neutrons were not sufficient to account for all the neutral current candidates, but the group was not confident enough to phrase their results in any terms but an upper bound on the ratio of neutral to charged events.[36] After the talk, Paschos called Musset to discuss with him the new results. Only a few weeks earlier, Paschos and Wolfenstein had published theoretical limits from the Weinberg model for precisely the process Musset and the hadronic group were studying. The theorists now had the result Musset had so wanted at that first meeting with Zumino in November 1971: a prediction that the NC/CC ratio would be above eighteen percent.[37] Such a fraction was just on the limit of

Figure 4. Transparency from Musset's talk at APS meeting in New York, January 1973. Number of events is plotted against longitudinal position in chamber. Object is to show the events are relatively constant in distribution as would be in neutrino interactions (but not in background neutron induced events).

X DISTRIBUTION ALONG BEAM



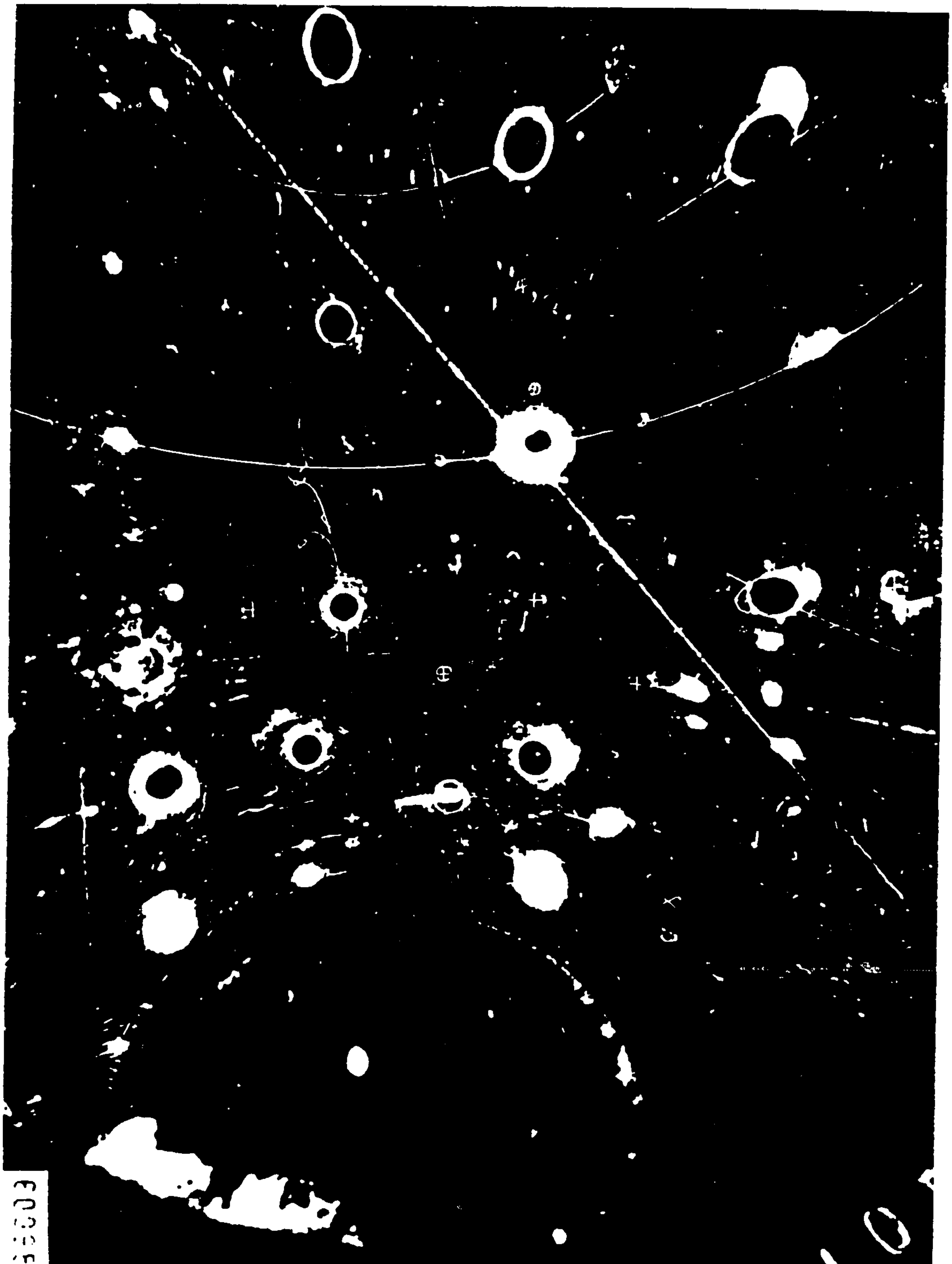
earlier bounds, and so was potentially compatible with what was known, yet was sufficiently large to be easily detectable.

Musset's excitement over the new theoretical results was augmented by another development he had heard about just a few days before leaving for the United States: in early January 1973, a single electron had been found during a routine rescan of some photographs at Aachen. (See Figure 5.) The Aachen electron satisfied all the criteria the Cundy group had imposed on it--it was well within the fiducial volume, making it almost certainly not due to a photon; it was of very high energy; there were no nearby events; and it was oriented in the direction of the neutrino beam.

Until January, Musset had felt he was working somewhat apart from the collaboration, mainly with the few people in the hadron group. When he returned from the States things had begun to change, partly as a result of the increasing numbers of hadronic muonless events and spatial distribution, and partly as a result of the single electron event. With positive results now coming both from the hadronic and leptonic groups, A. Rousset had the ammunition to request authorization for two more experiments, each with a million pictures.[38] The tone of his request reflects the more confident attitude of the neutrino group. He wrote:

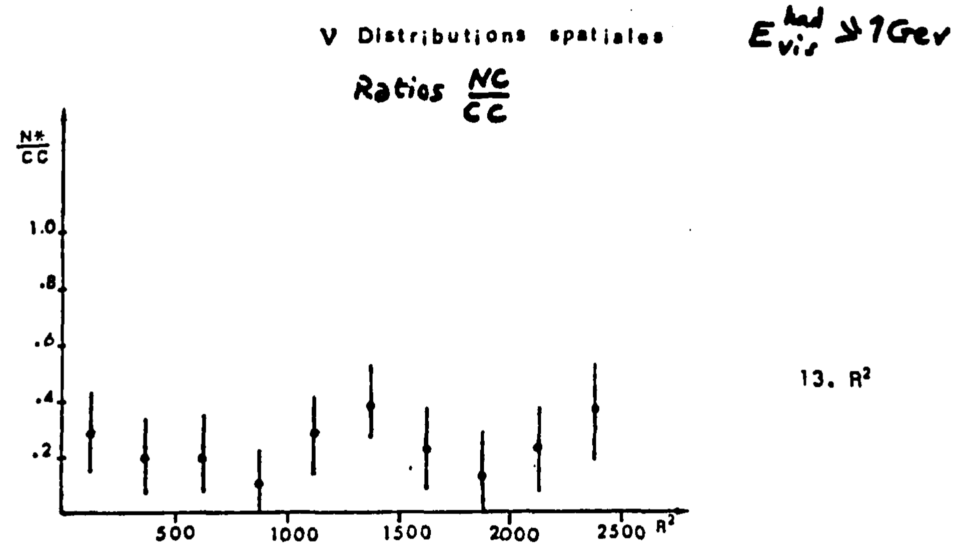
> The search for hadronic neutral weak currents in Gargamelle shows an appreciable amount of possible hadronic events (i.e. without charged lepton). These events have to be distinguished from neutron background. Severe cuts in the fiducial volume,

Figure 5. First single electron photograph from Gargamelle. Found at Aachen, January 1973. The electron's trajectory goes from right to left, beginning slightly below and to the left of the fiducial marker, ⊕.



> in energy, in angle are needed and the resulting statistics are then very small. A quick increase of statistics by a factor 2 to 3, possible with 2 to 3 weeks of running time, . . . would increase the significance of the results. In addition, one candidate of leptonic neutral current ($\nu_\mu + e \rightarrow \nu_\mu + e$) has been found in the present film and therefore we can hope to detect other events in the new film.[39]

With Lagarrigue and Rousset now both pushing hard for the neutral current search, this part of the program began to dominate all other neutrino work. By mid-March, Camerini and Musset had finished a new study of the old data tapes that had been prepared in December for the American Physical Society meeting. This time they put an even higher energy cut on the events in order to be even more sure the events were unaffected by the neutron background. An independent analysis of the same data was made by the Orsay group to check the results.[40] After the cut, they found that for neutrinos the ratio of neutral to charged events was 130NC/551CC (= .24), and for antineutrinos 83NC/191CC (= .43). (See Figure 6.) Immediately after presenting the numbers in the memorandum, Musset and Camerini added the results of the Weinberg model. It is clear that the data were perfectly compatible with the theory, but the crucial question remained: how many additional events needed to be subtracted because of background effects? Not surprisingly, the authors ended with a plea to "put priority on this study" of associated events which would help determine the neutron background.[41]



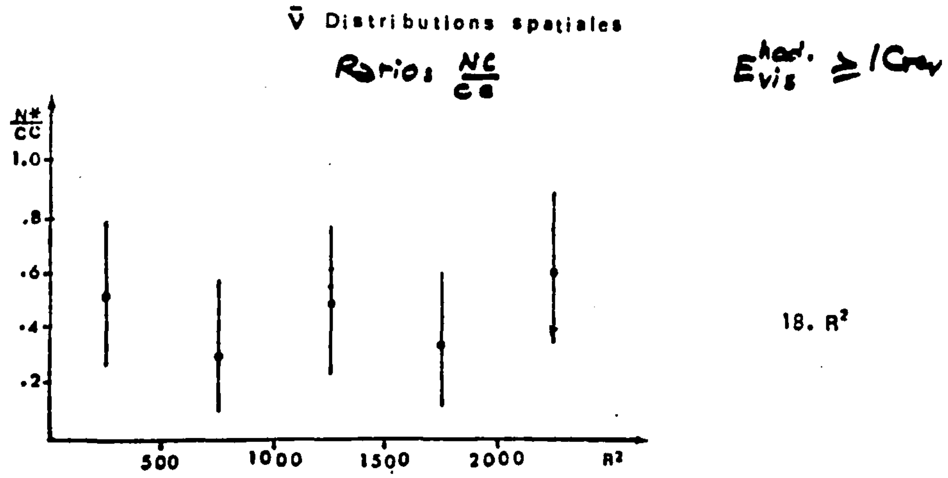


Figure 6. R = NC/CC plotted against position in chamber after energy cut off was made at 1 GeV. From P. Musset, CERN Technical Memorandum, 19 March 1973.

The help Musset needed was with the extraordinary amount of work required to study each neutral current candidate in detail. Huge enlargements of the appropriate photographs were made so that the group as a whole could judge their validity. The records from these meetings (see Figure 7) contain long lists of such judgments: "OUT possible μ," "OK one track badly measurable," "OUT cosmic [ray]," "OUT entering track," "OUT outside FV [fiducial volume] when best vertex," "K^0?," "OUT possible μ-kink," "OUT $E > 1$ GeV." These comments reflect the various tests events were subject to. The π's and p's could look like muons--this meant that to be conservative, possible charged events had to be discarded, as did cosmic ray events. Similarly, if there was evidence that a particle had entered with the beam in line with a vertex, the photograph was discarded as an "entering track." A "μ-kink" meant that the track suddenly bent--a possible sign that muon was really a pion or proton interacting with a nucleus. Other criteria such as the hadron energy cutoff at 1 GeV and the restriction of vertices to a fiducial volume helped statistically to ensure that neutron events would not be counted as neutrino events. All of these individual analyses, in addition, helped guarantee the same criteria were applied to neutral, charged, and associated events.

These small-scale debates occurred hundreds of times both at individual laboratories and at the larger neutrino collaboration meetings. Each decision added to what the participants hoped was a reasonably conservative estimate of the ratio, NC/CC. By April, the question of finding an upper limit for the ratio had been abandoned. The goal thereafter was to justify a number in accord with the Weinberg-Salam

Figure 7. Typical summary sheet of scanning meeting to evaluate the neutral current candidates and to make event categories standardize event criteria. From P. Musset, CERN TM, 14 April 1973.

Top half

TC-L/PA
PM/ju

17.4.1973

List of the NC events > 1 GeV
controlled at CERN on the 12 & 13 April 73

Total ν : 96
Total $\bar{\nu}$: 61

Event number	Classif.	Comments
AACHEN		ν 7 OK
399/276	OK	2 possible vertices
441/642	OUT	ν associated
556/077	OUT	only electrons possible ν_e
556/307	OUT	ν associated
556/543	(OUT)	probably < 1 GeV. Short track
570/097	OK	check if no possible μ. π^+ downstream ?
570/174	OUT	ν associated
570/253	OUT	< 1 GeV
570/399	OK	E badly known
707/244	OUT	ν_e associated
707/208	OUT	ν_e
714/252	OK	
742/205	OUT	entering track or ν event
749/449	OUT	E < 1 GeV possible μ
756/726	OK	
756/682	OUT	possible μ
756/376	OK	(> 1 GeV if π hypothesis)
763/150	OUT	possible μ
813/392	OK	
		$\bar{\nu}$ 11 OK
448/556	OK	one track badly measurable
462/295	OK	E > 1 GeV if π
483/665	OK	
525/090	OUT	entering track
598/247	OK	
605/033	OUT	possible μ
612/006	OUT	possible μ or e
612/648	OUT	ν associated
605/495	OK	E > 1 GeV ?
693/303	OK	
869/182	OUT	possible μ
883/389	OK	K° (6 γ) ?
897/131	OUT	E < 1 GeV
911/488	....	possible μ
918/451	OK	E > 1 GeV ?
932/115	OK	
932/369	OK	
932/512	OK	Dalitz

theory.

In the single electron search, confidence was also building up: the minutes from the meeting of 21 March 1973 begin with the remark that "There was general agreement that a paper should be published as soon as possible concerning the electron search and the one event found."[42] Even though all preliminary estimates indicated a very small background, some of the details needed to be cleaned up, such as scanning efficiency. Also, beta decay could yield a free electron which could be mistaken for a neutral current event if the proton had sufficiently low energy to remain undetected.

Simultaneously, several different groups worked to calculate the neutron background more rigorously: Vialle and Blum at Orsay using one kind of Monte Carlo, the CERN group using another.[43] Fry and Haidt very carefully calculated the probability of a neutron inducing a shower of other neutrons, thereby extending its effective range; Fiorini studied the problem of neutral kaons, while Pullia focussed on the attenuation of neutrons in the liquid. Lagarrigue made some rough calculations, and Rousset developed a thermodynamic analysis of the background treating the neutrons in equilibrium with the neutrino beam.[44]

Of the various approaches, Rousset's thermodynamic analysis of early spring 1973 was most persuasive for those members of the group who were not already convinced that neutral currents existed.[45] As in the other studies, the key quantity to estimate was the ratio of B (total number of neutron-induced NC events) to AS (the number of associated events which could be measured). Rousset's analysis was based on three simple equations:

(i) $N = B + AS$ (N = total rate of neutron interactions that look like NC events)

(ii) $N = \alpha N_\nu$ (N_ν is rate of neutrino events producing neutrons and α is proportion of neutrons that create events satisfying NC criteria, assuming an infinite length of liquid in which the neutron can interact)

(iii) $AS = \alpha N_\nu \langle p \rangle$ (AS = rate of associated events; $\langle p \rangle$ = probability of detecting a neutron interaction where neutron's origin is inside fiducial volume).

Therefore,

$$B/AS = (1/\langle p \rangle) - 1.$$

p (the distance a neutron travels before interacting) is a function of the spatial distribution of the neutron interactions. Suppose the spatial distribution is described by an exponential with characteristic interaction length λ. Then if L is the length from the neutrino interaction to the end of the fiducial volume, it follows that

$$p = 1 - \exp(-L/\lambda).$$

Thus,

$$\frac{B}{AS} = \frac{1}{\langle 1-\exp(-L/\lambda) \rangle} - 1$$

for the measured neutron interaction length. When other factors such as radial distribution of neutrino flux density of matter around the bubble chamber, neutron cascades, energy spectrum of neutrons, etc. are added, this result varied. But even taking these other effects into consideration, the neutron background could account for no more than 20% of the excess of NC events.[46]

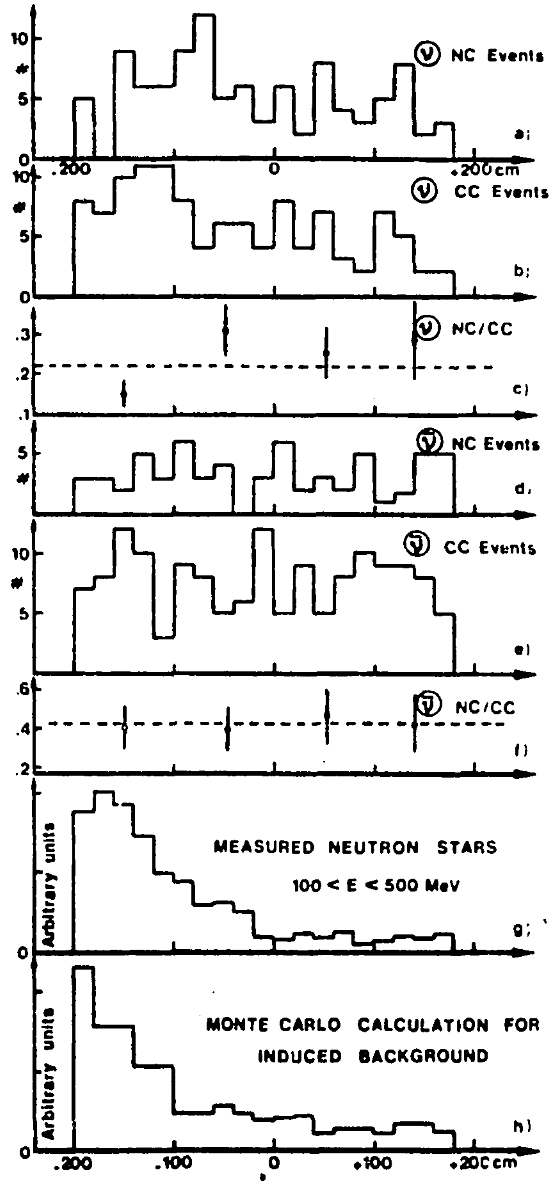
During these weeks, argument within the collaboration went back and forth as various members of the group suggested possible new sources of background, and others sought to demonstrate they could not be large enough to account for the excess of neutral current candidates. Vialle, for example, remembers Lagarrigue coming into his office practically every day with a new source of possible background.[47] Only days before Musset's seminar announcing the discovery of neutral currents, Pullia became very concerned about kaon regeneration, only to write to Musset a few days later that he had convinced himself it would not be a problem.[48] Thus, using a variety of approaches, techniques and approximations, the members of the collaboration persuaded themselves they were looking at a real effect.

The final argument, however, had nothing to do with the physics at Gargamelle. In early July, Carlo Rubbia, who also held a position at CERN, let it be known that the NAL group was close on Gargamelle's heels. According to many of the participants, this tipped the already tilting balance, and the decision to publish was made. Not everyone was entirely happy with the arguments presented in the final draft,[49] but they believed they had the background under control. On 19 July 1973, Musset gave a seminar at CERN announcing the discovery; four days later, on 23 July, the paper was sent to Physics Letters.[50]

In the Physics Letters article, the authors relied almost entirely on two arguments: (i) that various criteria such as spatial distribution and energy distribution were the same for CC and NC events and (ii) that the Monte Carlo predicted a neutron background significantly below the level of NC events found. Rousset's equilibrium argument and the associated studies of cascades, etc. were reduced to a single sentence. Their evidence was summarized in the following figures. (See Figure 8.) The group put their results in a conservative form, allowing that their data "could be attributed to neutral current induced reactions, other penetrating particles than ν_μ and ν_e, heavy leptons decaying mainly into hadrons, or by penetrating particles produced by neutrinos and in equilibrium with the ν beam." Nonetheless, the final sentence returned to the Weinberg model in concluding that their results would imply a Weinberg parameter $\sin^2\theta_W$ between .3 and .4.

By January of 1974, a more comprehensive summary of the work was prepared for Nuclear Physics B that included some of the values of B/AS generated by the Monte Carlo for a variety of values of neutron energy and angular distribution. Even under the worst case, the excess of muonless events to charged events was too large to be accounted for by the neutron background. In sum, they wrote, "The events behave as expected if they arise from neutral current processes induced by neutrinos and antineutrinos."[51]

Figure 8. Figures from first published paper on hadronic neutral currents by the CERN-Gargamelle collaboration. From F. J. Hasert et al., Physics Letters, 46b (1973), p. 139.



Distributions along the ν-beam axis. a) NC events in ν. b) CC events in ν (this distribution is based on a reference sample of ~ 1/4 of the total ν film). c) Ratio NC/CC in ν (normalized). d) NC in $\bar{\nu}$. e) CC events in $\bar{\nu}$. f) Ratio NC/CC in $\bar{\nu}$. g) Measured neutron stars with $100 < E < 500$ MeV having protons only. h) Computed distribution of the background events from the Monte-Carlo.

4. The First NAL Experiment 1A

Neutrino physics was one of the major justifications for building the National Accelerator Laboratory. The predecessor to experiment 1A (experiment 1), however, had its beginnings somewhat later, in the Summer Study program held by Fermilab in Aspen, Colorado. In 1969, while the accelerator itself was still under construction, many proposals were put forward for possible search programs for the W. One of the participants, A. K. Mann, presented a report at the school on the possibility of producing the W by means of a high energy neutrino source incident on a high-Z material, then detecting the particle's decay products using spark chambers between segments of earth. From his preliminary calculations, Mann argued, such a search could be effected up to a W mass of about 5 GeV. As it had been in the Schwartz and Yang and Lee program and the proposals drafted at CERN, the W-search was given the highest priority by many American experimentalists at Aspen.[52]

Partly as a result of his own work prepared for the conference and partly as a consequence of the other studies presented there, Mann advocated running a high-energy neutrino experiment at NAL.[53] But Mann was not the only physicist with his eye on the first neutrino experiment at NAL: it was clear from the start that whoever ran that first neutrino experiment would be in an excellent position to explore a region of energy high above that of all previous work. Thus, as Mann began to draft the proposal, it seemed to him likely that if the collaboration of Schwartz, Steinberger, and Lederman also submitted an application for the first neutrino experiment, they would present very stiff competition. Not only had these three experimentalists worked together before on the two-neutrino experiments, but the apparatus they had been using was quite similar to the spark chamber detector Mann hoped to build. In addition, J. Walker presented a proposal that remained in direct competition with experiment 1 until the final decision by the planning committee. To add more weight to his proposal, Mann turned to a younger physicist with whom he had been impressed, David Cline.

Cline, like Mann, brought with him much experience in experimental weak interaction pyysics. Ever since his Ph.D. he had been interested in the problem of neutral weak currents in kaon decay. As was mentioned earlier, such strangeness-changing neutral currents had quite severe upper limits placed on them by a variety of experiments. Almost all his career in physics Cline had been involved in these determinations using bubble chambers. For example, in 1964 with Camerini and Fry he had shown that the branching ratio of $K^+ \longrightarrow \pi^+ e^+ e^-$ was less than 10^{-6} of all K^+ decays.[54]

Much of Cline's careful work on the strangeness-changing neutral currents involved the identification of characteristic "signatures" of the various rare processes he was studying. The problem was to identify carefully as many unambiguous events as he could, thereby setting limits on the process. In the Camerini, Cline, and Fry paper, the authors had found three candidates for the $K^+ \longrightarrow \pi^+ e^+ e^-$ decay in the bubble chamber pictures and then had devoted a considerable part of the letter to an argument that two of the three must be due to a background process "faking" the neutral current events. Even the third, they concluded, was not "an unambiguous event and [we] shall consider it as an upper limit."[55] It was not, to use one of Cline's favorite phrases, the "gold-plated event" they were searching for. Still, it was good enough not to be discarded.

A few years later, after Cline and many others had conducted quite a variety of other experiments on strangeness-changing decays in a variety of channels, Cline reviewed the subject at the École Internationale de la Physique des Particules Élémentaires.[56] The tenor of the article was that the limits on neutral currents seemed to indicate such currents did not exist. Indeed, Cline summarized his review by declaring:

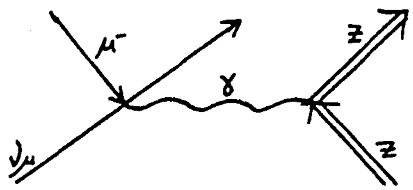
> . . . the crucial tests of such models [of weak interactions by Salam and Ward, Good, Michel, de Rafael, d'Espagnat, and Bludman] will probably come from experimental studies of lepton-lepton scattering which presently seem virtually impossible. Nevertheless, the successful explanation of the absence of neutral lepton couplings (and possibly of primative neutral hadron couplings) will undoubtedly be a very significant factor in the ultimate theory of the weak interactions.[57]

Cline was an important addition to the neutrino project, and, in December 1969, Cline and Mann drafted a more complete proposal for the experiment by elaborating Mann's earlier Summer School report.[58] Their goals were stated as three-fold. First, they wanted to measure the double differential cross section, $d^2\sigma/d(q^2)d(E_\nu - E_h)$ and the corresponding total cross section for $\nu_\mu + p \rightarrow \mu^- +$ anything.

The second goal of the experiment would then be the W-search. Assuming the mass of the W to be less than about 8 GeV/c^2, they would look for the particles through the reaction,

$$\nu_\mu + Z \rightarrow \mu^- + W^+ + Z$$

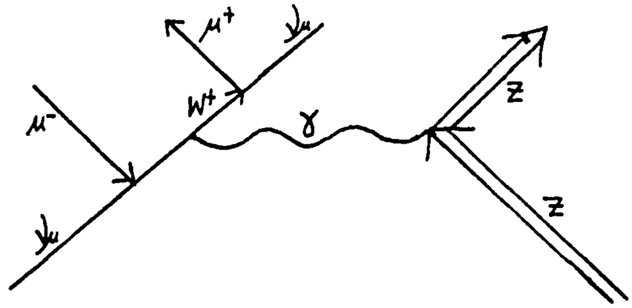
corresponding to the Feynman diagram:



If, on the other hand, the W was significantly more massive than 8 GeV/c^2, the authors hoped to look at the "point" interaction in which the decay of the W^+ would be seen. Nominally, the decay products of the W^+ would be $\mu^+ + \nu_\mu$, i.e.,

$$\nu_\mu + Z \longrightarrow Z + \mu^+ + \mu^- + \nu_\mu$$

corresponding to the Feynman diagram:



The proposed physics goals required a more sophisticated apparatus than a simple spark chamber, so Mann and Cline modified Mann's original detector in several ways. First, they proposed using liquid scintillator containers alternately placed between iron blocks to form a sampling ionization calorimeter. In this device, when hadrons hit the iron they caused further showers of charged particles. The cascade then caused light to be emitted as the charged particles collided with atoms in the scintillator. The light could then be collected and measured by phototubes. Second, 25 meters downstream from the calorimeter, they used blocks of iron alternately placed between spark chambers to determine the range (and therefore the remaining energy) of the muons. To determine the sign of the muons, the first section of the muon detector was to have been magnetized. Thus, by measuring the hadron and muon energies, the experiment in effect would determine the energy of the original neutrino, making possible the various cross section measurements proposed by the authors.[59]

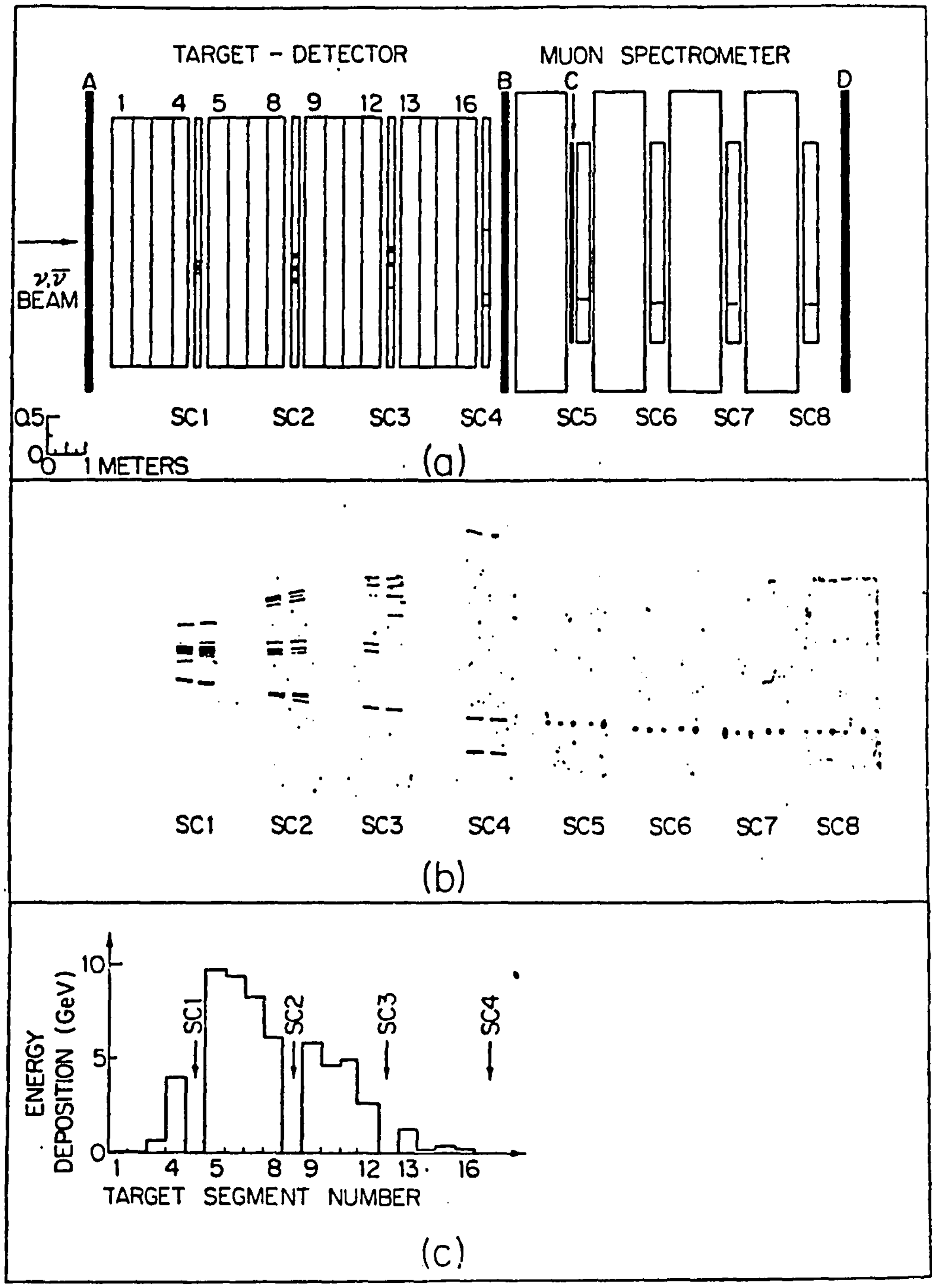
Mann nonetheless felt the apparatus was not yet well enough formulated to sway the planning committee at Nal so he and Cline turned to Carlo Rubbia (then at Harvard) whom Mann knew from a leave of absence Mann had taken at CERN.[60] Above all, Rubbia brought with him his experience designing and building large and sophisticated detectors.

During the previous several years, Rubbia had used large detectors to study the interference of K_S and K_L. This line of investigation which had begun as a test of CP conservation had yielded a wealth of new discoveries including Fitch and Cronin's discovery of CP violation.[61] One goal of the various collaborations in which Rubbia took part was to confirm the earlier results. Other objectives were to determine more precisely the $K_L - K_S$ mass difference and to obtain a better understanding of the empirical aspects of regeneration phenomena. All three of the principal collaborators of Experiment 1A thus came to the experiment with a strong background in weak interaction physics, though the kinds of experiments in which they had participated were quite diverse.

The three principals -- Mann, Cline, and Rubbia -- planned a meeting in the lobby of the JFK airport in late 1969. Before they parted ways again they had agreed to proceed with a joint proposal for the neutrino experiment.[62] When the proposal was finished, it had become the "Harvard-Pennsylvania-Wisconsin collaboration," and the four goals they set down were the same as the four main objectives of the CERN project being formulated on the other side of the Atlantic: 1. the W search (which they claimed could be undertaken up to ten GeV); 2, the point interaction $\nu_\mu + Z \rightarrow \nu_\mu + \mu^+ + \mu^- + Z$; and 3, the double differential and total cross sections for $\nu_\mu + p \rightarrow \mu^-$ + anything. In addition the group now identified the probing of hadronic structure in the deep inelastic region (large $E_\nu - E_h$) as one of their primary interests. For only since the SLAC deep inelastic results and the corresponding speculation on their theoretical origin, had hadronic structure become a major concern. As it had at CERN, the parton model took its place beside the W search as a major goal of the experiment.

To accomplish these goals, Rubbia and his collaborators redesigned the earlier two-stage detector in several important ways (see Figure 9). The calorimeter was redesigned to be totally active, that is, all the

Figure 9. 9a is the design of experiment 1A. First stage consists of liquid scintillators alternately placed between spark chambers. This serves both as a target and calorimeter to measure total energy deposited by interaction. Second stage consists of iron slabs alternately placed between spark chambers; the whole second stage is in a magnetic field and thus serves as a muon spectrometer. 9b is a typical charged current event (muon present). 9c is the calorimeter record of event in 9b. Figure (of 9a, 9b, and 9c) from A. Benvenuti et al., Physical Review Letters, 32 (1973), p. 801.



TARGET - DETECTOR
MUON SPECTROMETER
A
B
C
D
1
4
5
8
9
12
13
16
$\nu,\bar{\nu}$ BEAM
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC8
0.5
0
0
1 METERS
(a)
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC8
(b)
ENERGY DEPOSITION (GeV)
10
5
0
SC1
SC2
SC3
SC4
1
4
5
8
9
12
13
16
TARGET SEGMENT NUMBER
(c)

energy deposited in the mineral-oil based liquid scintillator would be collected by phototubes. Between the scintillator containers were placed spark chambers to record both hadron and muon tracks. In addition, counters A, B, C, and D could be used to trigger the recording devices selectively. For instance, they would be triggered only when no charged particles entered the device through A in time with a hadron shower in the calorimeter. The second stage of the detector was improved as well. Instead of determining the energy of the muons by their range through blocks of iron, the group installed large magnetized iron blocks, which served to measure the muon's momentum by its deflection. From the total energy deposited in the calorimeter and the muon energy as determined in the spectrometer, the neutrino energy could be calculated.[63]

Thus, if the physics goals resembled those of CERN, the experiment itself certainly did not. The idea behind the design of the apparatus remained as a two-stage analysis of the neutrino interactions: calorimeter and muon spectrometer. By combining these two detectors, the 1A group would record more information than one would in a simple spark chamber experiment, and therefore the group would be able to compete favorably with the bubble chamber neutrino groups. In addition, the spark chamber-calorimeter had two other important advantages. First, spark chambers could be built much larger than bubble chambers, giving a ratio of 10 (100 tons versus 10 tons) in the target mass. Further, E1A would operate at 10 times the energy of CERN-Gargamelle (20 GeV versus 2 GeV), providing yet another factor of ten in the expected rate of neutrino interaction. Thus, on the order of 100 times the

Gargamelle rate per day could be expected at NAL. A second advantage of the spark chamber was that by being active, it could exclude events in which an unseen neutron interacted with nuclei creating a shower of charged particles.

The competition between the two types of detectors thus reflected a deep experimental conundrum: bubble chambers provided great detail on particle momenta and identification, but they are passive devices requiring vast amounts of film and running time to locate rare events. Spark chambers, by contrast, normally offered less detail in the event analysis, but are active, recording event information only when specific logic circuits are fired, and provided a much higher rate of interactions. E1A was centered around a detector which was designed to try to bridge this gap, if only partially. As we will see, this dilemma, pitting particle identification against high statistics, played a crucial role in the subsequent neutral current search.

Neutral currents, it should be added, figured but little in the Harvard-Wisconin-Pennsylvania proposal. They are not mentioned in the primary physics objectives. But more important, the design of the apparatus was such that even in principle (in its original form), the experiment was not capable of a neutral current search. The reason neutral currents could not be found is that the logic circuits would have an event recorded only if a muon penetrated into the muon spectrometer; unfortunately, neutral current events were characterized precisely by having no muon. This feature of the trigger had been borrowed (along with much else) from the Schwartz, Steinberger, and Lederman experiment,[64] where such a trigger was crucial to eliminate extraneous events where no muon was produced. Finally, as in the CERN proposals, even where neutral currents were mentioned (in the context of dimuon production), there was no mention of Weinberg-Salam theory at all, and no quantitative prediction was discussed of the order of magnitude effect to be expected.

During the winter and spring of 1970, plans for the experiment advanced, and in the summer of 1970 Cline, Mann, and Rubbia published an article describing another channel through which they could use their apparatus to detect the decaying W's. This article, "Detection of the Weak Intermediate Boson Through Its Hadronic Decay Modes,"[65] again focussed on the search for intermediate bosons in the energy range of 5-10 GeV/c^2. In print, the HWPF did not, however, discuss neutral currents (or the weak interaction in general) in relation to the Weinberg-Salam model before 1972.

In January, February, and March of 1972, E-1A and E-21 (Barish et al.) began skirmishing over who would actually be the first to run a neutrino experiment on the new beam. After several exchanges of letters and meetings with the director, Wilson let E-1A proceed as the first experiment.

Meanwhile, t'Hooft's renormalization proof and Weinberg and Salam's theory reopened interest in the gauge-theoretical unification ideas. Once again, NAL and CERN continued to move in parallel. Whereas in Switzerland, Zumino had come to speak to the experimentalists about the consequences of the gauge theories, in America, after the renormalization proofs, Weinberg began calculating some experimental consequences of his theory--calculations which until then had not seemed worth undertaking. Some nine years later, Weinberg recalled,

> . . . Now we had a comprehensive quantum field theory . . . the weak and electromagnetic interactions that was physically and mathematically satisfactory in the same sense as quantum electrodynamics--a theory that treated photons and intermediate vector bosons on the same footing, that was based on an exact symmetry principle, and that allowed one to carry calculations to any desired degree of accuracy. To test this theory, it had now become urgent to settle the question of neutral currents.[66]

Weinberg published calculations of the cross sections to be expected for neutral current production. In addition, from MIT, he called Rubbia at Harvard to tell him how important it was for the NAL group to search for the expected muonless events.[67] Rubbia recalled that,

> Steven Weinberg was the one, who, with rare insistence . . . was chasing me and many other people [to do the neutral current search]. I learned all these things [about gauge theories and neutral currents] from him directly. I remember I was down in the old cyclotron at 4 Oxford Street. He called me up--in the beginning I thought, my God, what [is] he asking me to think? [Then] I realized how beautiful things were.[68]

Soon, the E1A collaboration decided to do the search; it fit in with some of their earlier interests and seemed possible without extensive modification of the apparatus. It also added yet another reason for the steering committee at NAL to choose E1A, as they were quick to point out to the director of the laboratory, Bob Wilson:

> There has recently been increasing awareness of the need of more sensitive searches for neutral weak currents and neutral weak intermediate bosons. The existence of a neutral weak current or a neutral weak propagator would cast additional light on the connection between weak and electromagnetic interactions. As the center of mass energy, $S^{\frac{1}{2}}$, available to experiments increases, and GS moves closer in magnitude to α, the possibility of finding such a connection becomes more realistic. We might now stand in a position analogous to that of Oersted, Ampère and Faraday 150 years ago as they attempted to elicit the connection between electricity and magnetism.
>
> We have observed, along with others, that a sensitive test of a recent, possibly renormalizable, theory of weak interactions may be made through comparison of the observed rates for the processes $\nu_\mu (\bar{\nu}_\mu) + N \longrightarrow \nu_\mu (\bar{\nu}_\mu) +$ anything and $\nu_\mu (\bar{\nu}_\mu) + N \longrightarrow \mu^{\mp}$ + anything, where N is a

> nucleon. Different models allow for some leeway in the expected value of the ratio $\sigma(\nu_\mu)/\sigma(\mu)$ but a value $\lesssim 0.01$ would be quite difficult to accommodate in that theory.[69]

The immediate experimental necessity was to install a trigger on the calorimeter that would fire if <u>either</u> the hadron energy was above a certain minimum in the calorimeter <u>or</u> a muon passed into the muon spectrometer. Rubbia commented later that he had been in favor of putting the trigger into the experiment, "not because I had decided it [beforehand] , but because Steve Weinberg gave me a good reason for it."[70] The actual construction of the trigger was the first independent task that Larry Sulak, then a young assistant professor at Harvard, undertook on the project. Aside from the immediate problem of putting together the electronics, it engaged Sulak full-time in the problem of the neutral current search.

Data from the experiment came in painfully slowly. The beam was on for a few days near Thanksgiving 1972, then again for a short time near Christmas. Between the two runs, the energy trigger yielded some 150 events to be examined; these were first assessed by the Wisconsin group with Sulak flying out to help. Soon, however, the data were brought to Harvard which became, for the first part of the experiment (up to August 1973), the focal point for the neutral current search.

Almost as soon as the energy trigger was installed, pictures began to show up without muons (pictures like the ones reproduced at the beginning of this essay). Much later these were taken to be photographs of a process including weak neutral currents. But at least Mann saw them quite differently at the time; he later commented in an interview:

> You can say, well, we came to the conclusion immediately that we had seen weak neutral currents. But you'd be surprised, that was the last conclusion we came to. Our first conclusion was that we were making some mistake and that these muons were somehow escaping the apparatus or being missed by us in some way and that no effect of that magnitude could exist.[71]

It must be reemphasized that Cline, and Mann, independently,[72] had conducted precise measurements to show that neutral currents in kaon decay did not exist in some channels above one part in a million. Only later was it accepted that charm suppressed neutral current decays if they were strangeness-changing, but did not affect the strangeness-conserving processes considered by ElA and Gargamelle. At the time, neither the 1A group nor anyone else sought to draw a radical division between strangeness-changing and strangeness-conserving decays. It was therefore natural when dealing with a new machine for the experimentalists to suspect that some error was producing the ratio of over 30% muonless events to events with muons.

Consequently, during the spring of 1973, Mann and Cline were concerned principally with understanding the physics of charged current events and various other projects originally set out as goals for experiment 1A. Their reasoning was that the charged current events would yield information about the properties of the detector as well as about the charged current events themselves. With so many aspects of the beam and detector still untested, this seemed a necessary prerequisite for the study of any new physics including neutral currents. Culminating these first efforts was a paper submitted by the group to Physical Review Letters entitled, "Early Observation of Neutrino and Antineutrino Events at High Energies."[73]

Meanwhile, during the spring of 1973, Sulak began the analysis of the films brought back from the experiment. Several undergraduates assisted him, and the small group remained in contact with the larger group and with Rubbia, who was travelling back and forth to CERN.[74]

From the computer tape, Sulak determined the frame numbers on the film of events where more than a minimum cutoff amount of energy was deposited in the calorimeter. Then, frame by frame, in the fourth floor room in Lyman Laboratory at Harvard, he and the undergraduates studied the photographs in a high accuracy film projector, sorting muonless from charged current events, and measuring the properties of both.

The problem of escaping muons was their overriding concern. (See Figure 10.) Since any individual "muonless" event might have a muon escaping detection by exiting from the calorimeter at a large angle, it was necessary to work out a computer-simulated model for wide-angle muons. By comparing the number of muons expected not to reach the muon spectrometer with the number of measured muonless events, they could determine if there was a statistically significant excess of neutral

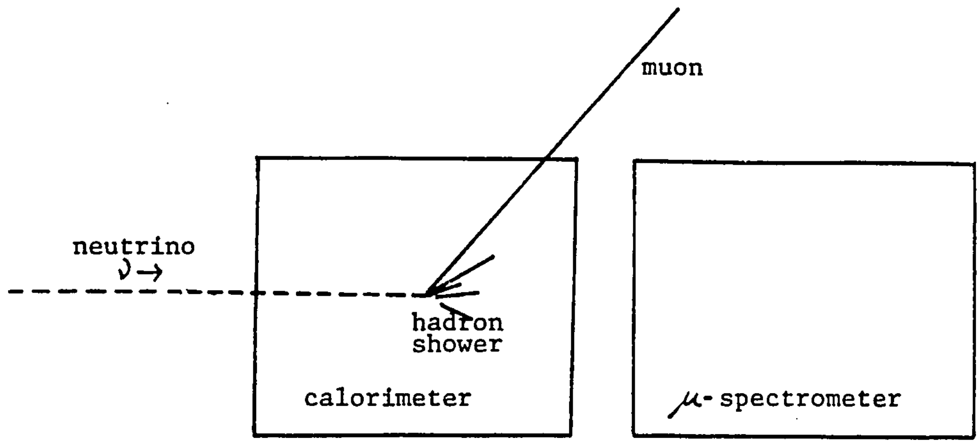


Figure 10. Wide-angle muon escapes detection in the μ-spectrometer thereby making the charged current event look like a neutral current event.

candidates. The first version of the Monte Carlo program distributed the muons, using an angular distribution given by the parton model that had emerged from the SLAC experiment discussed earlier.

When the Monte Carlo results were ready and compared with the first batch of photographs, it became clear that there was an excess of muonless events. After correction, the ratio, R_1 was found to be:

$$R = NC/CC = .41 \pm .08.$$ [75]

During June and July, Sulak and the undergraduates prepared a Physical Review Letter. Meanwhile, at NAL, Bill Ford, an Assistant Professor at Pennsylvania, and others began to work on the 400 GeV data. These had been obtained later and so were not included in the first muonless event analysis. Ford began later than the Harvard group and so when the paper was finally ready in late July, only about half as much data existed at 400 GeV as at 300 GeV, but they seemed statistically in accord with the lower energy results. Sulak then brought the manuscript to Mann (who was sick in bed with back problems), and Mann, Ford, and Cline agreed the paper should be submitted for publication.

All of this work in the late spring of 1972 was done knowing that CERN was accumulating evidence on the weak neutral currents, since Rubbia was commuting regularly between CERN and the U.S. and others from the CERN group occasionally visited NAL. In mid-July, Rubbia, independently, wrote a letter to Lagarrigue telling him of the recent NAL work:

> I have heard from several people at CERN that your neutrino experiment in Gargamelle in addition to the beautiful electron event has now a growing evidence for neutral currents. We have observed at NAL approximately one hundred unambiguous events of this type and we are in the phase of final write-up of the results. In view of the significance of the result I am addressing to you this note in order to know if announcing our result we should mention the existence of your work on the hadronic processes (and if so in which form). In this case I hope you will take a similar attitude toward our work.[76]

Lagarrigue declined Rubbia's offer the next day, suggesting that the announcements be made independently without mentioning the other's results, adding that the CERN announcement would be made in twenty-four hours, on 19 July.[77]

Upon returning to the United States, Rubbia helped make final revisions of the NAL paper which was widely distributed as a preprint in late July and August. After Rubbia's departure from the U.S. on 27 July, Sulak finished the draft of the paper and brought it by hand on 3 August to George Trigg, the editor of Physical Review Letters. This draft, meanwhile, had been seen by a fair number of theorists and experimentalists, and based on their comments the group made some corrections. First, the theoretical angular distribution that had been used to generate events in the Monte Carlo was replaced by an empirical one, based on the muon distribution in the last few chambers of the calorimeter. Second, more data were included from analyses at Madison and Philadelphia.[78] However, when Ford's new data were compiled, they showed a significantly lower value of R, especially in the first six segments of the calorimeter where they now found R equal to $.06 \pm .16$, thus only half of one standard deviation from zero. The average value of R from all the 300 and 400 GeV data was therefore revised from $.42 \pm .08$ to $.20 \pm .09$.

A technical digression is necessary here. The authors wrote in their table 1 of the revised paper that they had a 5.2 standard deviation effect, a remark which caused a great deal of controversy and confusion. The justification for this number was based on the following statistical distinction. There are two ways to measure the statistical significance of the value of R determined by the group. (1) The question can be asked, "How well is the value of R known?" for which the answer depends on the uncertainty of R, that is, on the $\pm .09$. (2) The question can be asked, "Given the assumption that the pre-Weinberg-Salam theory of weak interactions is valid (i.e., that there are only higher order effects simulating neutral currents), what is the probability that one would find a value of R = .29?" The answer to this second question depends not at all on the uncertainty in R, but only on the distribution of R's to be expected from the old theory of weak interaction. In other words, method 2 gave the probability of the effect not being a random fluctuation from the predictions of the old physics. This latter approach characterized the overall point of the Harvard paper. They did not want to stress the particular value of R, but only that neutral currents existed.[79]

On the basis of their statistical evidence for the effect, Rubbia and Sulak began to prepare for the summer conferences at Aix-en-Provence and Bonn, where they would announce their findings. In late August, Sulak brought the data over to Europe (where Rubbia had remained since leaving the U.S.) and they, along with Jim Pilcher and Don Reeder, went to Bonn for the International Conference on Electron and Photon Interactions at High Energies. There, they tried unsuccessfully to convince the organizers, especially George Myatt, to allow them to present their data at a plenary session.[80] Myatt, a member of the CERN-Gargamelle collaboration, had come to speak on neutral currents; he did, however, agree to include a paragraph on the NAL results.

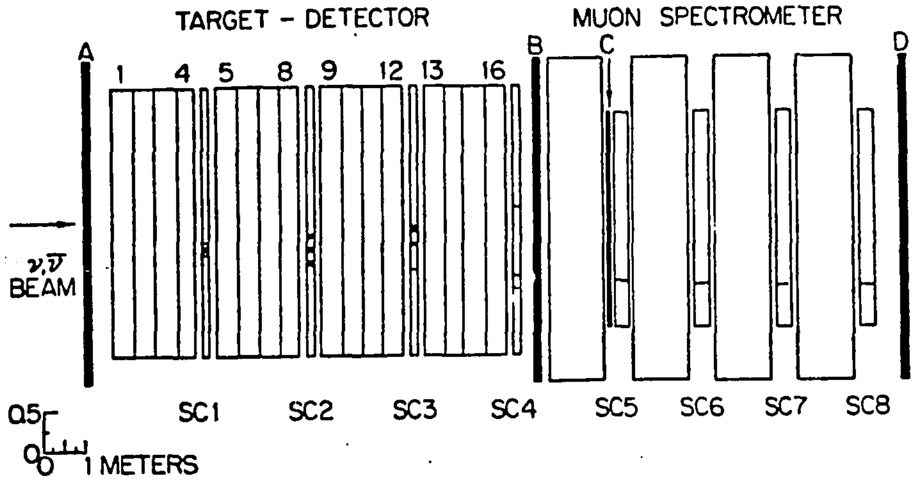
After the talk, Myatt was asked how these results of CERN and NAL could be reconciled with the low limits on strangeness-changing neutral currents in K and Σ decays. "That," he responded, "is a major obstacle to the Weinberg-type theories." This exchange is important because it makes it clear that even after the existence of neutral currents was being established, the charm hypothesis was not widely accepted, even among the participants in the neutral current search. Despite this, Sulak, Rubbia, and the CERN group had convinced themselves that they had observed neutral currents in the strangeness-conserving processes they were studying. Indeed, by late summer, after the conferences, it seemed for the Harvard group that the experiment was over.

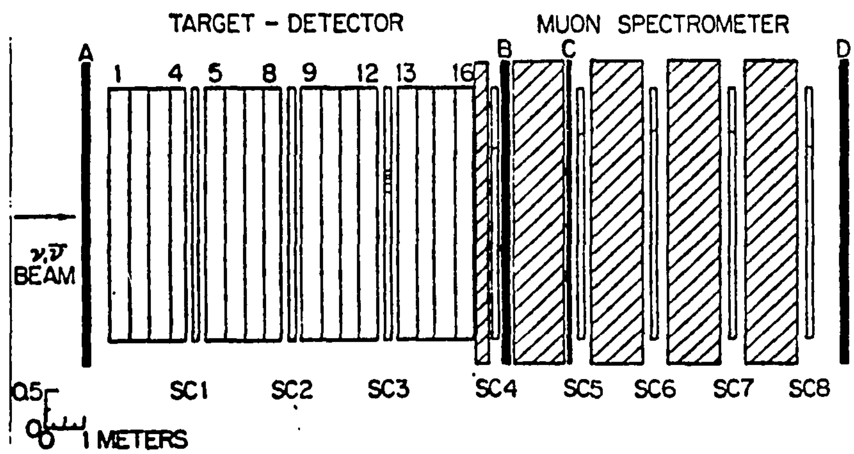
5. The Second NAL Experiment 1A

At NAL, however, it was just beginning. Four circumstances contributed to a certain distrust Cline and Mann felt about the paper submitted to Physical Review Letters. First, the 400 GeV data reduced at Madison indicated a very low ratio of neutral to charged currents. Second, Cline at least came to the experiment having repeatedly set extremely low limits on neutral current processes in the kaon decays. Not unreasonably, he expected in the summer of 1973 to place yet another low upper bound on the neutral currents. Given the uncertainty in the use of the new apparatus in addition to the wide angle muon problem, it was natural that he sought a further check on the new results. Finally, Mann felt that the whole experiment could be redone rapidly in a much improved way. Consequently, the full attention of Cline, Mann, and the others at NAL was, as a result, devoted to the rearrangment of the detector. For the moment, believing the conference reports to be a sufficient description of their work, the paper was put on the back burner and the referees' criticism was not immediately answered.

The main improvement Cline and Mann sought to make was to move a counter in the muon spectrometer closer to the calorimeter to catch more of the wide angle muons. (See Figure 11.) In addition, they replaced the spark chambers with larger ones which also improved the angular acceptance of the muon spectrometer. The price they had to pay for these changes at the time did not seem high; they were forced to introduce a new, 13" thick steel shield to separate the calorimeter from spark chamber 4, which then could serve as a wide angle muon

Figure 11. (Top) Old apparatus as described in Figure 9. (Bottom) New arrangement uses spark chamber 4 (which previously was part of stage 1) to capture muons at wider angles than previously possible. To filter hadrons out, a thin iron plate (13") is placed in front of spark chamber 4. From B. Aubert et al., Physical Review Letters, 32 (1974), p. 1455.


TARGET - DETECTOR
MUON SPECTROMETER
A
B
C
D
1
4
5
8
9
12
13
16
ν,ν̄
BEAM
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC8
0.5
0
0
1 METERS


TARGET - DETECTOR
MUON SPECTROMETER
A
B
C
D
1
4
5
8
9
12
13
16
ν,ν̄
BEAM
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC8
0.5
0
0
1 METERS

detector. This shield plus the downstream sections of the calorimeter would presumably stop the hadrons formed in the upstream part of the calorimeter from penetrating into the spectrometer and thus impersonating muons. Previously, this function had been served by a much thicker (4') iron slab that had come before the first counter in the spectrometer. But now with the steel slab wedged before the last spark chamber, the slab needed to be thinner to allow the last spark chamber to be close enough to the calorimeter to catch the wide angle muons. Cline commented on the change in a memorandum shortly after the first test run of the new apparatus on 28 September:

> The new iron placed behind the calorimeter is very effective in reducing the hadron penetration to . . . [spark chamber 4]. Some small number of events do show penetration but the fraction is very likely less than 20% . . . . More study of the data is needed to make this a reliable conclusion.[81]

Unfortunately, though it was not to be understood for several months in a quantitative way, the shield was not thick enough to be very effective in reducing hadron penetration. This was a crucial problem. For if the hadrons penetrated through the iron, even if no muon emerged from the vertex, the event would be recorded as a charged current event. (See Figure 12.) By not compensating adequately for the punch through, the neutral current signal would seem to plummet towards zero. The reason precise predictions could not be calculated for the hadron punch through is related to the reason the Gargamelle group was having

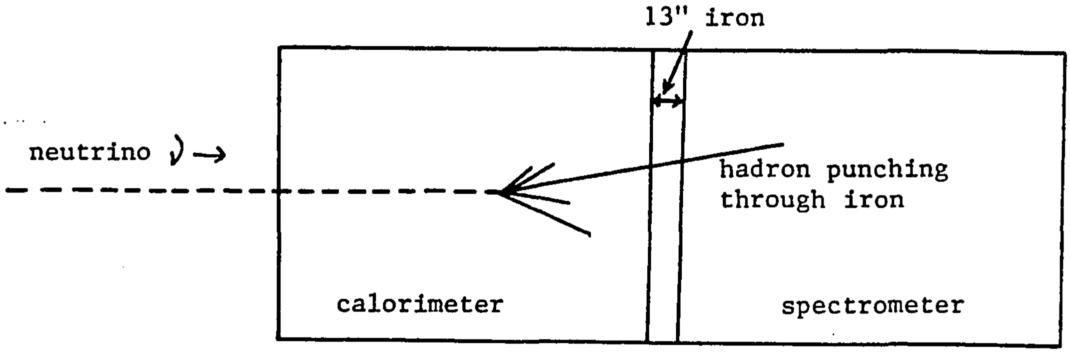


Figure 12. Punch through. Hadrons penetrate into muon spectrometer.

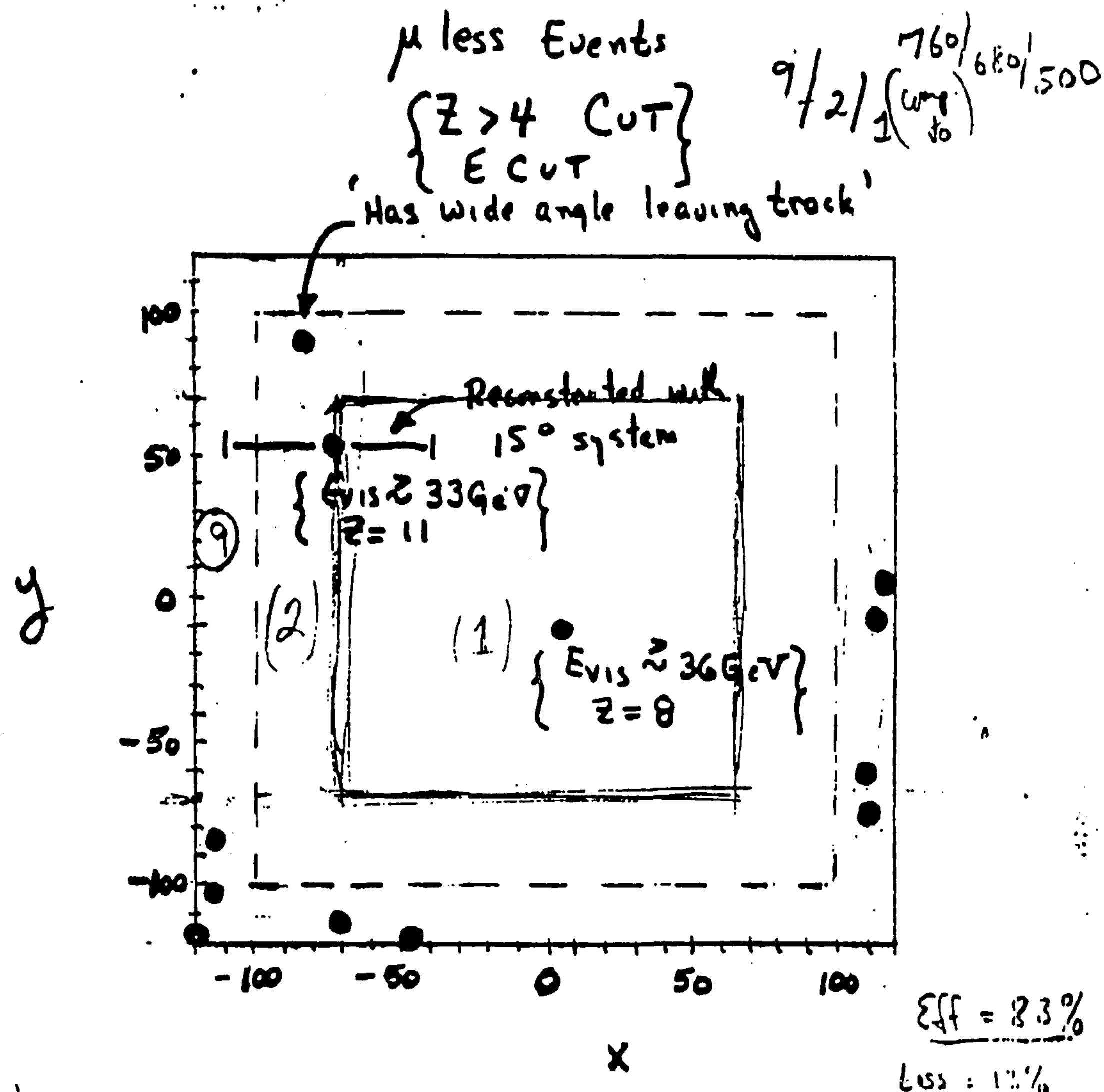
such a hard time calculating the neutron interaction length: both problems involved the passage through strongly interacting particles. This is a much more difficult problem than the well-understood electromagnetic interactions involved for instance in a muon's passage through matter. Compounding the problem was the absence of good data on the energy and momentum distribution of the hadrons being produced. This was the first observation of high energy neutrino reactions; and the composition of the reaction products had not been studied at all. Since punch through had not been a dominant problem during the earlier experiment, it was not at first realized that the thinner shield made it a serious one now.

In part, this was because the NAL group at this point was still looking for single unambiguous events, the kind of "gold-plated events" that Cline had successfully used before in his bubble chamber work to set very low limits on neutral current processes in kaon decay. It was therefore natural for him to continue to look in E1A for the same type of argument. In the same memorandum, Cline took the vertex reconstruction and other information from the data tapes to examine a single event, dead center in the fiducial volume, which had survived both position and energy cuts. (See Figure 13.) "It is amusing," Cline wrote,

> to investigate how improbable the central (x,y) event is . . . . (The other two events are too close to the edge of the fiducial region to be gold plated.) . . . we expect to find . . . 1 events. Thus, unfortunately this event is not improbable and we have not found a gold plated event.[82]

One corollary of this style of work (in which one searched for "shining examples") was that Cline was not especially confident in the statistical approach on which the initial paper was based.[83] Such computer simulations seemed to him much too vulnerable to errors in fixing the various parameters such as the characteristics of the neutrino beam and the muon angular distribution. Mann too felt doubtful about the earlier Monte Carlos, and sought to recheck the angular distribution of muons (for charged events) to large angles. This would

Figure 13. Candidate for a "Gold Plated Event." D. Cline, NAL, Technical Memorandum, 1 October 1973.


μ less Events
{Z > 4 CUT
E CUT}
9/2/1 (wng. ν0) 760/680/500
'Has wide angle leaving track'
Reconstructed with 15° system
{Evis ≈ 33GeV
Z = 11}
{Evis ≈ 36GeV
Z = 8}
(9)
(2)
(1)
y
x
100
50
0
-50
-100
-100
-50
0
50
100
Eff = 83%
Loss : 17%

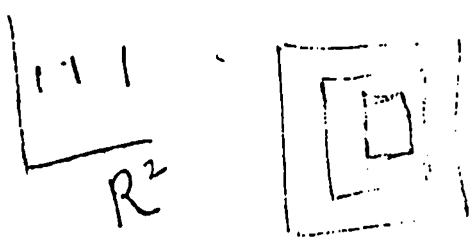

R²

FIG 4

ensure, he felt, that the corrections for wide angle muons were being made properly for the muonless events.[84]

Cline's doubts about the existence of neutral currents were expressed a few days later in a technical memorandum suggesting it would be interesting to look for muonless events that might arise from the production and decay of intermediate vector bosons--a possibility completely incompatible with the Weinberg-Salam theory at the energies they were using. The following day, 11 October, Cline sent out the first preliminary indications that experiment 1A no longer was giving results compatible with CERN's publication.[85] The calculations were crude, using two crucial numbers: Reeder had calculated an 83% muon detection efficiency, and Ling had estimated a 13% hadron punch through. This last number was less than half of what it was eventually found to be, and had the effect of lowering radically the number of calculated excess muonless events. Since more pions were penetrating through the steel than they thought, many real muonless events were being counted as charged current events.

For a variety of reasons, this error persisted for some time before a rigorous analysis was undertaken. First, in the old experiment, hadron punch through had not been a problem because of the thicker iron shield. Second, the physics of hadron interactions in iron at high energies was not especially well understood or measured at the time. Third, the energy distribution of the pions was not well known. Fourth, the group was under enormous pressure to present a result. Finally, the Madison group were now finding what they thought they would find:

that the muonless events were an artifact of the apparatus' geometry.

On 11 October, Cline circulated a memorandum placing a 90% confidence limit on R of .07, and a 99% confidence upper limit of .21. "Taken at face value," he concluded, "these results are inconsistent with the CERN measurement of R = .28 $\pm$ .03 for a mixed beam [of neutrinos and antineutrinos]. Clearly, it could still be that we did that one in 100 experiments or something else is wrong."[86] Something else was wrong, but it would take the group two more months to be sure what it was.

The pressure, meanwhile, was building up. Cline recalls getting less and less sleep as the project was stepped up to provide a definite answer to the neutral current question.[87] On 16 October 1973, another memorandum was distributed, also by Cline:

> (i) Because of the importance of the neutral current question, the fact that we have extended our necks previously on the subject and that other groups around the world are moving fast to check our results and the CERN result I propose that a rapid, unified analysis of the μless events be carried out early in November at NAL . . . .
>
> (ii) The schedule of our run has changed with the laboratory now inserting running time for E21 at the end of November. I suspect that this time will be used for a μless search since they are likely submitting a proposal for this experi-

> ment in the next week or two. Again this proves the need for us to move fast in our analysis and to settle the question before others get to it.[88]

That same day, 16 October 1973, the referee reports from the Harvard paper were sent back from Physical Review Letters to Sulak.[89] Both referees agreed that corrections were necessary to clarify the wide angle muon problem. Both also criticized the way the statistics had been handled, claiming that Sulak and Rubbia's technique for assessing the statistical significance of the data was not sufficiently conservative. Essentially, both referees wanted the authors to base their conclusions on the uncertainty in R rather than on the probability that E1A's value of R (.29) was compatible with pre-Weinberg-Salam physics. Indeed, both referees recommended against publication of the paper until their objections were satisfied. However, Rubbia was out of the country, Sulak was at Harvard, and Cline and Mann were preoccupied with the revised experiment. The referees' comments as expressed in the report were therefore not answered until several months later.

The referees were not the only ones with doubts about the NAL procedure. From Europe, Bernard Aubert, who had worked with the Gargamelle group until August and then transferred to the NAL group, reported that he was spending his time defending the NAL experiment to the neutrino physicists there.[90] According to Aubert, the NAL's "lack of credibility . . . comes mainly from the fact that [the European physicists] do not know how well we measure E_μ and E_h and [they] believe that we guess more than measure the $[E_\mu/E_h]$ uncertainty."

By mid-November, Mann and Cline were convinced that the newer results definitely failed to give evidence of neutral currents; Rubbia concurred. Mann then drafted a letter to this effect for the Physical Review Letters, which was intended to replace the earlier paper that had been allowed to sit at the offices of Physical Review Letters pending the outcome of the revised experiment. Though the "No Neutral Currents" article was never actually submitted, it represented a good summary of the state of opinion at that time. (See Figure 14.)[91] The abstract read in part as follows:

> The ratio of muonless events to events with muons is observed to be .05 ± .05 for the specific case of an enriched antineutrino beam. This appears to be in disagreement with recent observations made at CERN and with the predictions of the Weinberg model.[92]

There was some division in the NAL group over the question of how and when the new results should be released. Mann felt the group should wait before discussing them. Rubbia and Cline at different times discussed the current situation with people outside E1A.[93] When Rubbia went back to CERN in December 1973, he spoke with a variety of people, including Musset, Lagarrigue, A. Rousset, Jentschke, and others.[94] By this time, the CERN group had, of course, already published their result that neutral currents did exist; naturally, they were somewhat distressed.

Jentschke, then director general of CERN, convoked a meeting of the Gargamelle group to cross-examine them on the experiment; he was

Figure 14. Draft of Physical Review Letter, asserting E1A shows no evidence for neutral currents at Weinberg-Salam level. This paper was never published.

% 929 Rough Draft (1)

Search for Neutrino Induced Events Without a Muon in the Final State*

B. Aubert‡, A. Benvenuti, D. Cline, W. T. Ford, R. Imlay
T. Y. Ling, A. K. Mann, F. Messing, R. L. Piccioni,
J. Pilcher*‡, D. D. Reeder, C. Rubbia, R. Stefanski and L. Sulak

Is the background data needed?

Department of Physics*
Harvard University
Cambridge, Massachusetts 02138

Department of Physics*
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Department of Physics*
University of Wisconsin
Madison, Wisconsin 53706

National Accelerator Laboratory+
Batavia, Illinois 60510

Abstract

A comprehensive search for neutrino induced muonless events has been carried out using a liquid scintillator calorimeter - magnetic spectrometer exposed to various neutrino beams produced at the National Accelerator Laboratory. The ratio of muonless events to events with muons is observed to be 0.05 ± 0.05 for the specific case of an enriched antineutrino beam. This appears to be in disagreement with recent observations made at CERN and with the predictions of the Weinberg model.

afraid that CERN would be publicly embarrassed by the forthcoming American announcement. The Gargamelle group, however, would not back down. Still, they were shaken.[95] Musset circulated a memorandum advising the groups to deemphasize the Weinberg theory and to redouble effort on the study of the associated events. The memorandum began:

> Dear Friends,
>
> After our last neutral current meeting, all of you have probably heard rumours about new results in the WB beam at Batavia with a slightly changed apparatus (muon counter after 1 foot of iron) and a focusing horn for an $\bar{\nu}$ run. The efficiency for μ detection is better than previously and the result is an apparent lack of neutral current type events.
>
> In the near future, we can expect to be heavily questioned about the reliability of our experiment.
>
> Independently from these new rumours, it is much more important to know if neutral current type events can be simulated by a trivial background such as neutrino induced neutrons, than to measure accurately a $\sin^2\theta_W$.[96]

At the National Accelerator Laboratory, not only Imlay, but Aubert, Ling, and Sulak were working on the punch through problem nearly full-time. Preliminary results indicated that the punch through was higher than at first thought, but that the calculations were still not as high as they would be eventually.[97]

When the "No Neutral Currents" paper was completed in draft form, Mann composed a letter to the CERN group informing them of the result. Mann, Cline, and Rubbia signed it. But before sending it, the authors (Mann, Cline, Rubbia, and Reeder) consulted with R. R. Wilson, then director of Fermilab. Wilson advised them to wait somewhat longer until the experiment was complete before announcing the result. Though the letter was as a result never sent, Rubbia brought an unsigned copy to its intended recipient, Lagarrigue. From his office it was duplicated and several members of the CERN collaboration were familiar with it. (See Figure 15.) The authors concluded in their letter that the corrected ratio of muonless to charge current events was:

$$R = .02 \begin{smallmatrix} + .05 \\ - .03 \end{smallmatrix} ,$$

in other words, <u>statistically indistinguishable from zero</u>.[98] The trade of punch through for a better angular acceptance had exacted a higher price than they knew.

For until early November only the simplest attempt had been made to measure punch through. A pencilled comment by Cline in the margin of the "No Neutral Currents" draft said, "R.I. Imlay, do this [punch through] calculation." Adding to the uncertainty was the fact that different approaches to measuring the punch through probability at first yielded different results. For instance, Sulak measured the ratio of the number of events where many sparks appeared in the first spark chamber after the thin iron plate, to the number of events where only one spark was found there. Nominally, such a ratio should approximate the percentage of hadrons penetrating through the iron plate along with

Figure 15. Letter from D. Cline, A. K. Mann, D. D. Reeder, and C. Rubbia to A. Lagarrigue. This letter was never sent, but an unsigned duplicate was brought by Rubbia to Lagarrigue and was then seen by many members of the CERN collaboration.

November 13, 1973

Professor A. Lagarrigue, Director
Linear Accelerator Laboratory
University of Paris - SUD
Centre D'Orsay
Batiment 200
91405 Orsay
France

Dear Professor Lagarrigue:

We write to inform you of the preliminary result of our recent experiment to search for neutrino interactions without final state muons. As you know, our apparatus was modified to provide a much larger detection efficiency for muons relative to the apparatus that was used in our earlier search for muonless events. We also improved our ability to locate accurately vertices of observed neutrino interactions, and lowered the threshold on the total energy of the hadrons in the final state.

From about one half of the data obtained in our recent run, we find the raw ratio $R_{raw} = 0.18 \pm 0.03$. We estimate the muon detection efficiency of the apparatus for the enriched antineutrino beam that was used in this experiment to be approximately 0.85. Taking into account small backgrounds produced by incident neutrons and by ν_e in the incident beam, the corrected ratio is $R_{corr} = 0.02 \begin{smallmatrix} + 0.05 \\ - 0.03 \end{smallmatrix}$, where the error includes an estimate of the uncertainty in the calculated detection efficiency. We are continuing to process the remainder of the data and to improve our understanding of the experiment.

We have written a paper intended for Physical Review Letters which will soon be submitted. A copy will, of course, be sent to you but for obvious reasons we wanted to convey our result informally to you before its publication.

With kindest regards

Yours sincerely,

D. Cline D Cline

A. K. Mann A K Mann

D. D. Reeder

AKM/rs

C. Rubbia Carlo

the muons. One problem with this method was that individual sparks often did not show up very well; another was that the two stereo cameras gave divergent results:

multiple sparks/single sparks:	
x view:	15%
y view:	30%[99]

Another approach Sulak mentioned was to plot the number of muonless events as a function of longitudinal distance along the machine. This plot indicated a rapid decrease, while the number of charged plus muonless events remained constant. (The total number presumably was the total number of neutrino events which should remain constant throughout the detector.)

Here too the interpretation remained ambiguous. On the one hand, one could say that there were few muonless events downstream because pions were punching through. On the other hand, one could say there were more muonless events upstream only because the wide angle muons were escaping. Sulak left "the conclusions to the reader!" Since he was an advocate of the original Monte Carlo, he hoped to convince the others that the punch through was causing the rapid decrease.[100]

However, Mann and Cline used this data to conclude that there were no neutral currents, since they were mostly concerned with wide angle muons. Thus, when they extrapolated the ratio of neutral to charged events to the last segments before the iron plate, they found

$$R \text{ (corrected)} = .057 \pm .053,$$

a result in good accord with the one written up in the "No Muonless

Events" Physical Review Letter.[101]

Summing up these various punch through studies and a further one by Imlay,[101] based on pion penetration measured by earlier experiments, Cline gave a talk at NAL on E1A's latest results.[102] When the values for the geometric efficiency (ε) and punch through (ε_p) were inserted, Cline arrived at a corrected R between .05 and .15 which was both below the CERN data and the Weinberg prediction. Cline concluded (see Figure 16):

> 1. R' Very likely too small to be consistent with Weinberg model and lower bounds deduced by Paschos and Wolfenstein for this model--also CERN data, if due to Weinberg model--Energy dependence is still a loophole.[103]

This last remark referred to the possibility then being entertained by some members of the collaboration that the discrepancy between CERN and NAL might be attributed to the difference in their neutrino energies. (In retrospect, we know this not to be the case.) The conclusion continued:

> 2. R' = 0.29 $\pm$ 0.09 Suggested by first E1A experiment is not confirmed in the present experiment--uncertainty in the (x,y) vertex reconstruction in that experiment was perhaps the trouble--there are still loopholes however![104]

By adding a new camera, the stereo photographs now yielded a more accurate location of the vertex; this was not as it turned out of great

Figure 16. Concluding transparency from D. Cline's talk at NAL, 6 December 1973.

TENTATIVE
Conclusions [redacted]

1. R' Very likely too small to be consistent with Weinberg model and lower bounds deduced by Paschos and Wolfenstein for this model – also CERN data, if due to Weinberg model – Energy dependence is still a loop hole

$R' > 0.28$ CERN at $E_\nu \sim 2$ GeV

2. $R' = 0.29 \pm 0.09$ Suggested by first E1A experiment is [redacted] not confirmed in the present experiment – uncertainty in the (x,y) vertex reconstruction in that experiment was perhaps the trouble – there are still loop holes however!

3. There are events in the present experiment that don't appear to have a visible μ track – could be due to spark chamber ineff. – but seems unlikely – needs further study

4. Some μ less events appear to have showers that are suggestive of predominate electromagnetic effects – perhaps due to ν_e's

importance, but at the time was thought to be a possible explanation for the old results.

It is important to remember that throughout this time in September, October, November, and December, the group was under a great deal of pressure to announce their findings. This pressure came not just from the other experimentalists, but also from the theorists who were getting informal progress reports from various participants. As Mann later recounted:

> As the results began to emerge we were being pressed harder and harder for some kind of decisive answer from people. It is very hard to communicate to you how [things were], when you are in the center of the stage at a time like that, particularly in high energy physics where you do not quite have control over your own destiny. You have to work with collaborators, with the lab, with the director, with the program committee and with all the people who do the chores that allow the experiment to be done. You're being leaned on over and over again to produce whether you're ready to produce or not.[105]

During this period, each of the participants was struggling to integrate the various calculations and measurements; each had to convince himself of the reality or artificiality of the effect. Every measurement and calculation had its own weaknesses and strengths best known by the individual or subgroup involved. As a result, both of pressure from outside the collaboration and of new evidence from within the group, opinions were changing. On 13 December, Cline sent out a memorandum with a new tone:

> Three pieces of evidence now in hand point to the distinct possibility that a μless signal of order 10% is showing up in the data. At present I don't see how to make these effects go away.[106]

The three pieces of evidence were first, the Monte Carlo now yielded an R = .1; second, the spatial distribution of the events looked as if it had been caused by true neutrino events. But what I take to be the most convincing for Cline was the third reason he offered: among twenty neutral current candidates, five "had no hint of wide angle tracks."

> These events were in the center of the detector and the μ angles would have to be at least 200 - 300 mr and with the result that the μ track will be <u>well separated from the rest of the shower</u>. The separation should help increase the sparking efficiency. It seems unlikely that the chamber efficiency goes to 25% for such modest angles. . . . This is certainly consistent with a true μless signal of R' . . .=.08.[107]

This was the kind of argument Cline liked: a small selection of events, clean of possible edge effects and with an analysis that did not require resorting to Monte Carlo techniques.

About the same time, Mann convinced himself that he agreed: the signal would not go away.[108] Over the course of December and January 1974, Mann examined the data and photographs again, applying various selection criteria to be sure no simple error would account for most of the muonless events. Just as in the CERN meetings, the events were scanned and rescanned, energy was remeasured, fiducial regions redefined, pictures were rechecked for through muons, and so on. In a final, internal report of 26 January 1974, Mann argued his new position:

> . . . it appears that our scanning criteria and fiducial region cuts, (x,y) 120 and 5 Z 12, do in fact eliminate most questionable events. Of 13 "N" [neutral current candidate] events in runs 328 to 332 in the final sample, 8 of them are "good to look at," as the attached reproductions indicate.[109]

In addition, a new Monte Carlo calculation was almost ready. Five days after Cline's "ten percent" memorandum, Aubert, Ling, and Imlay completed a detailed and rigorous study of the punch through which could then be used to generate an accurate assessment of the background. When this was done, the signal was raised to a 12 to 15% level.[110] The second version of the 1A experiment thus neared completion, and after several meetings during January and February, it was decided to publish the original Harvard paper with the comment that additional work had confirmed the earlier findings.

By the end of February, the 1A group had essentially finished two separate experiments. Not only was the beam changed, the geometry shifted, the spark chambers replaced, the background different, but the participants in the two experiments were not all the same. The style of experimentation of the two subgroups was different, their expectations were not the same, and the evidence that finally convinced them that neutral currents were real was different.

In mid-March of 1974, the collaboration finished a paper on the revised experiment entitled "Further Observation of Muonless Neutrino-Induced Inelastic Interactions," which they sent to Physical Review Letters.[111] By then the new evidence was presented in its most convincing form, concisely summarized in the following nine figures. Most of the symbols used in the figure captions have been defined earlier. The others are: $\bar{A}EBC$ = anti coincidence of counter A with energy trigger and coincidence of counters B and C; SC4 = Spark Chamber 4 which is now serving as the first muon detector; E_μ = geometric efficiency of muon detector; R = ratio NC/CC; (on diagrams of Figure 17 below) μ_1 = muon detection by SC4 and counter B; μ_1' = muon detection by SC4 alone; μ_2 = muon detection by SC5 (the first muon detector of the old experiment).

In the future, it will undoubtedly be for these and similar diagrams that the work of E1A will be remembered. Indeed, with this paper, the first chapter of the discovery of weak neutral currents drew to an end. Further experiments were performed at many laboratories all over

VOLUME 32, NUMBER 25 PHYSICAL REVIEW LETTERS 24 JUNE 1974

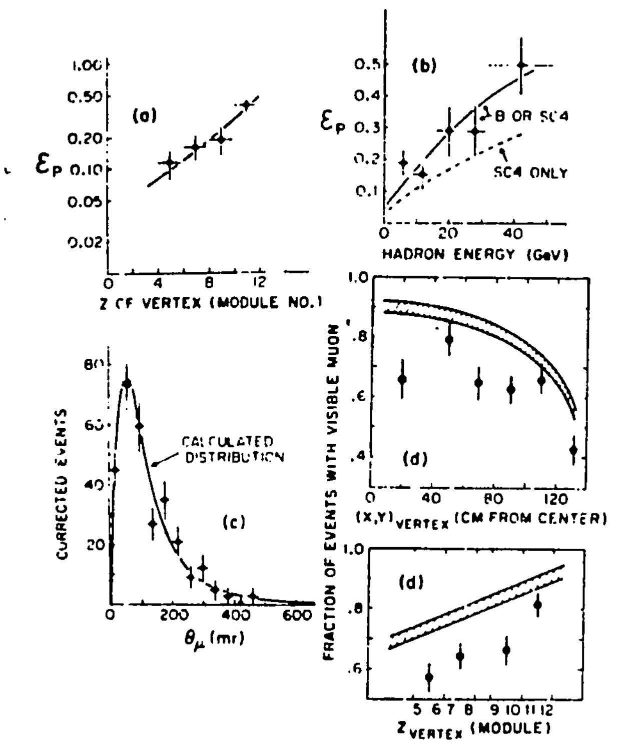


FIG. 3. (a) The measured punch-through probability of hadrons accompanying $\bar{A}EBC$ events (for all hadron energies) as a function of z, and the expected shape of the distribution. (b) The measured punch-through probability (for z between 5 and 12) as a function of E_H, compared with the expected variation. (c) The corrected muon angular distribution measured in SC4 compared with the predicted distribution. (d) Comparison of the observed fraction of events with a muon for the μ_1' identifier (SC4 alone) and ϵ_μ as functions of transverse position and z position. The cross-hatching indicates the uncertainty in ϵ_μ arising from the statistics of the data in (c).

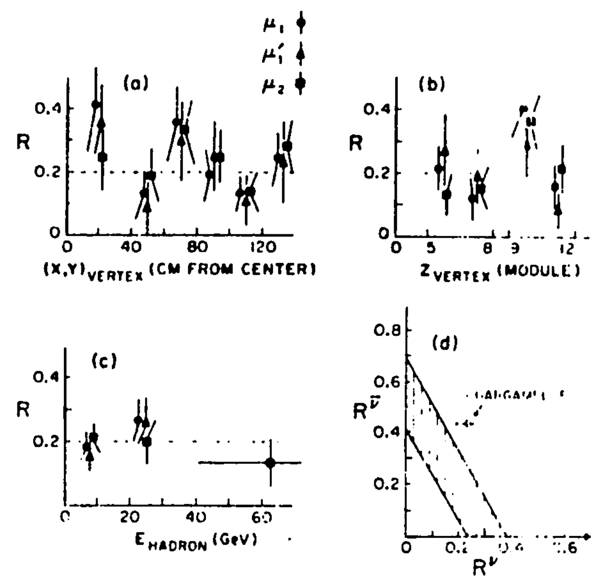


FIG. 4. (a) R obtained from three different muon identifiers as a function of the transverse distance from the center of the calorimeter. (b) The z variation of R obtained using three different muon identifiers. (c) The E_H variation of R from three muon identifiers. (d) The allowed region of R^ν and $R^{\bar\nu}$ from this experiment compared with R^ν and $R^{\bar\nu}$ obtained in the CERN measurement (Ref. 2).

Figure 17. Evidence for neutral currents from E1A's second neutral current publication. B. Aubert et al., Physical Review Letters, 32 (1974), p. 1456.

the world to determine the space-time and isotopic spin structure of the currents, but the existence of the currents themselves seemed to be assured. Twice over, the NAL collaboration had had to struggle through the long, difficult, and frustrating task of separating artifact from reality.

6. Conclusion: The End of Experiments

The principal goals of experiment 1A and Gargamelle evolved in parallel from a search for the intermediate vector boson to a study of the parton model, and only gradually became an investigation of weak neutral currents. At each step on the way, experiment and theory interacted to advance together.

There are similarities as well between the two groups as each moved towards the discovery of neutral currents. Within the broad framework of the participants' interest in weak interaction physics and neutral currents, each collaboration fell into two subgroups. One group within each experiment came to the experiment with the experience of having looked for and not found neutral currents. In E1A, this experience involved putting extremely low upper bounds on the strangeness-changing decays; in Gargamelle, it included setting some (incorrectly low) limits on strangeness-conserving processes. From all sides the evidence suggested a virtually complete suppression of neutral current processes. Furthermore, at the time practically no one felt a strong reason to expect different results in strangeness-changing and strangeness-conserving events, even though the explanation of the difference (in terms of charm) had appeared in print. In Gargamelle, there was the additional factor that all those having participated in the earlier CERN bubble chamber work knew very well how difficult it was to extract a signal from the neutral background.

Similarly, in both E1A and Gargamelle, there was another group that, from 1971 onwards, put priority on the hadronic neutral current search. In E1A, part of this interest came from Weinberg who was pushing some of the participants to look for the hadronic neutral currents at the 20% level. For the CERN group, Zumino, Paschos, and Wolfenstein played the analogous role.

By thus persuading some members of the two collaborations that a search at the 20% level was worthwhile--in fact, urgent from the theorists' perspective--the Weinberg-Salam theory exerted the first stage of its influence. The immediate result at 1A was the installation of a trigger that would fire when a certain amount of hadron energy was deposited, even if no muon emerged, thus providing at least the possibility of a neutral weak current search. At CERN, the main consequence was the establishment of what at first was an informal group of physicists (later a formal subgroup of the collaboration) to scan, measure, and select the hadronic neutral current events.

In a certain limited sense, the neutral currents were "there" from the start: both NAL and CERN had photographs they would eventually present as evidence for weak neutral currents. The real work of the experiments, however, was for the collaborators to convince themselves that the photographs were photographs of something real and not an artifact induced by the apparatus or environment. What followed was almost a year and a half of a seemingly endless list of internal debates over the tracks and sparks, the acceptance, the efficiency, the neutron background, the muon spectrum, the neutrino flux, the beam purity, the through muons, the fiducial volumes, the cosmic rays, the neutral kaons, and the statistical significance of the results. Many of these sub-experiments required a commitment of weeks, sometimes months; each helped to expand the circle of participants convinced that the effect existed above the background. On the whole, these subexperiments took place in the domain of established physics, meaning within established experimental techniques and theoretical ideas. Delineating the background thus formed the second interaction of theory and experiment.

These two conceptual moments, the design of the experiment (both of goals and apparatus), and the decision of when to stop the experiment, both need to be studied to understand the history of an experiment. Traditionally, textbook and even most historical accounts have left out the latter, limiting themselves to a description of what the experiment was intended to determine, and then discussing the results eventually found. In doing so, they leave out the flesh and blood of the experiment.

One might call this second stage of experiment-theory interaction a process of validation. In the experiments studied here, this process took place in several different ways. During the original E1A experiment, with the analysis based at Harvard, the validation was accomplished primarily by a combination of variation of the fiducial volume to show that the percentage of neutral currents remained relatively constant and by comparing the experimental number of neutral current events to the number predicted by the Monte-Carlo. When the experiment was redone by the Wisconsin group, the validation process focussed more on the punch through and other machine characteristics than on the Monte Carlo. In Cline's work especially, one sees the reflection of a style of work developed earlier in the study of rare events. In bubble chamber work, during the early stages of the revised experiment, he was searching for some "golden events" by which he could validate the neutral currents. At CERN, the validation process for the hadronic neutral currents took place primarily through the many different neutron background analyses. Musset was persuaded by the relative number and spatial distribution of associated events; some other collaborators were persuaded by the thermodynamic analysis, yet others by the Monte Carlos. Still others remained skeptical until the problems of the neutron cascade and kaon regeneration were fully understood.

Only gradually were the various indidivual arguments transformed into the kind of evidence finally assembled for publication. Little by little, the conclusion was built up out of the many steps necessary to assess the background. Certainly no one moment can be pointed to either in ElA or in Gargamelle that could be called the time of the discovery.

Finally, I would like to suggest that it is by studying how the process of validation has changed over time that we can understand the nature of the change in experiment since the turn of the century. Before the time of high energy physics, experiments were conducted primarily either by one or two physicists. The change in the scale of the apparatus has necessitated much larger groups. This has had several effects. First, validation is accomplished by small subgroups working on specific problems instead of one or two people conducting variations on the main experiment. Second, since the apparatus itself is very likely to have no close duplicate, many of its properties have to be understood as part of the experiment. In earlier, smaller-scale experiments, such investigations would have been conducted as separate experiments, as was possible with some of the more standardized equipment of cloud chamber experiments in the mid-thirties.

Consequently, in many ways the large experiment now subsumes into its own internal dynamics the processes which previously took place in the scientific community as a whole. This is visible not only in the large-scale repetition of ElA within the same experiment, but also in the multiplicity of subgroups working separately and partly independently to determine the punch through in ElA or the neutron background in Gargamelle. It is visible too in the role of the internal publication of reports and memoranda.

If I could finish with one suggestion: in the history of experiment, past and contemporary, we must focus attention on the process by which the experiment was ended as well as by how it began. For the decision that an effect is real brings together the social dynamics, the theoretical assumptions, the experimental technique, and the individual styles of research. When looked at in this way, contemporary experiments suggest that the logic of discovery and logic of justification lose some of their distinct identities.

We need a richer descriptive vocabulary to describe experimentation in a way that will account for the many intermediate steps between the often very subjective working hypotheses of various participants, and the logically or empirically based argument that eventually finds its way to publication. Such a vocabulary would be able to depict the degrees of persuasive force that evidence has as it begins to accumulate from diverse considerations. In the process of developing an account like this, we will come to understand how evidence is transformed (as in the case of the first muonless event photographs) from curiosities to convincing evidence.

NOTES

[1]A. K. Mann, "Present Status of Inelastic and Elastic Hadronic Weak Neutral Currents," in Unification of Elementary Forces and Gauge Theories, papers presented at Ben Lee Memorial International Conference on Parity Nonconservation, Weak Neutral Currents and Gauge Theories, Fermilab, October 20-22, 1977. A. Rousset, "Neutral Currents," paper presented at the International Conference on Neutrino Physics and Astrophysics, Philadelphia, 26-28 April 1974. Printed in Neutrino-1974, C. Baltay, ed. (New York: American Institute of Physics, 1974), pp. 141-165. C. C. Llewyn Smith, "Unified Models of Weak and Electromagnetic Interactions," in Proceedings of the 6th International Symposium on Electron and Photon Interactions at High Energies (Amsterdam: North-Holland Press, 1974), pp. 449-465. C. Baltay, "Neutrino Interactions II: Neutral Current Interactions and Charm Production," in Proceedings of the 19th International Conference in High Energy Physics, Tokyo, 1978, pp. 882-903. G. Myatt, "Neutral Currents," in Proceedings of the 6th International Symposium on Electron and Photon Interactions at High Energies, Bonn, 27-31 August, 1973 (Amsterdam: North-Holland Press, 1974), pp. 389-406. J. E. Kim et al., "A Theoretical and Experimental Review of the Weak Neutral Current: A Determination of Its Structure and Limits on Deviations from the Minimal SU(2) X U(1) Electroweak Theory," Reviews of Modern Physics, 53 (1981), pp. 211-252. D. Cline and W. F. Fry, "Neutrino Scattering and New-Particle Production,"

Annual Review of Nuclear Science, 27 (1977), pp. 209-278, esp. pp. 233-247.

[2]S. L. Glashow and H. Georgi, "Unified Weak and Electromagnetic Interactions Without Neutral Currents," Physical Review Letters, 28 (1972), pp. 1494-1497.

[3]M. Schwartz, "Feasibility of Using High Energy Neutrinos To Study the Weak Interactions," Physical Review Letters, 4 (1960), pp. 306-307. T. D. Lee and C. N. Yang, "Theoretical Discussions on Possible High-Energy Neutrino Experiments," Physical Review Letters, 4 (1960), pp. 307-311.

[4]G. Danby et al., "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos," Physical Review Letters, 9 (1962), pp. 36-44.

[5]See, for instance, the introduction to P. K. Kabir, ed., The Development of Weak Interaction Theory (New York: Gordon and Breach, 1963). See also, Allan Franklin, "The Discovery and Nondiscovery of Parity Nonconservation," Studies in the History and Philosophy of Science, 10 (1979), pp. 201-257.

[6]T. D. Lee and C. N. Yang, "Implications of the Intermediate Boson Basis of the Weak Interactions: Existence of a Quartet of Intermediate Bosons and Their Dual Isotopic Spin Transformation Properties," Physical Review, 119 (1960), pp. 1410-1419.

[7]Ibid., p. 1410.

[8]"Projet de Chambre à bulles liquides lourds de 17m3," typescript report. Also author's taped interview with P. Musset, 26 November 1980.

[9]Weinberg-Salam theory puts the mass of the W at about 80 GeV, though it still has not been found.

[10]"Draft Gargamelle Proposal," typescript, 4 February 1970.

[11]M. Briedenbach, et al., "Observed Behavior of Highly Inelastic Electron-Proton Scattering," Physical Review Letters, 23 (1969), pp. 935-939.

[12]R. Feynman, High Energy Collisions, Third International Conference, State University of New York, Stony Brook, 1969, C. N. Yang, J. A. Cole, M. Good, R. Hwa, and J. Lee-Franzini, eds. (New York: Gordon and Breach, 1969). See also, J. D. Bjorken and E. A. Paschos, "Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon," Physical Review, 185 (1969), pp. 1975-1982.

[13]D. H. Perkins, "Draft Gargamelle Proposal," p. 1.

[14]Author's taped interview with J. P. Vialle, 28 November, 1980.

[15]S. Weinberg, "A Model of Leptons," Physical Review Letters, 19 (1967), pp. 1264-1266. A. Salam, "Weak and Electromagnetic Interactions," in Elementary Particle Theory, ed., N. Svartholm (Stockholm: Almquist and Forlag, 1968).

[16]G. t'Hooft, "Renormalization of Massless Yang-Mills Fields," Nuclear Physics B, 33 (1971), pp. 173-199.

[17]E. A. Paschos and L. Wolfenstein, "Tests for Neutral Currents in Neutrino Interactions," *Physical Review D*, *7* (1973), pp. 91-95. A. Pais and S. B. Treiman, "Neutral-Current Effects in a Class of Gauge Field Theories," *Physical Review D*, *6* (1972), pp. 2700-2703. S. Weinberg, "Effects of a Neutral Intermediate Boson in Semileptonic Processes," *Physical Review D*, *5* (1972), pp. 1412-1417. G. t'Hooft, "Prediction for Neutrino-Electron Cross-Sections in Weinberg's Model of Weak Interactions," *Physics Letters*, *37B* (1971), pp. 195-196.

[18]For a co-citation analysis of the changing interest in Weinberg-Salam theory after t'Hooft's paper and through the discovery of neutral currents, see D. Sullivan, et al., "Understanding Rapid Theoretical Change in Particle Physics: A Month-By-Month Co-Citation Analysis," *Scientometrics*, *2* (1980), pp. 309-318.

[19]One relevant experiment was described by W. Lee, "Experimental Limit on the Neutral Current in the Semileptonic Process," *Physics Letters*, *40B* (1972), pp. 423-425, and by B. W. Lee, "The Process $\nu_\mu + p \rightarrow \nu_\mu + p + \pi^0$ in Weinberg's "Model of Weak Interactions," ibid., pp. 420-422. On p. 422, B. W. Lee wrote that their experiment taken with the electron neutrino results "rules out the existence of the neutral current predicted by Weinberg's model of weak interactions." Musset and other members of the CERN collaboration were suspicious of this result from the start because of charge exchange corrections which had not been made by Lee and Lee in their analysis. See D. H. Perkins, "Neutrino Interactions," *XVI International Conference on High Energy Physics* (Chicago, 1972), esp. pp. 205ff.

[20] D. H. Perkins presenting work of H. H. Bingham, et al., "CERN Neutrino Experiment--Preliminary Results," Proceedings of the Sienna International Conference on Elementary Particles 30 September - 5 October 1963, Volume 1 (Bologna: Societá Italiana di Fisica, 1963), pp. 555-584. In an appended conclusion, J. S. Bell, J. Løvseth, and M. Veltman, "CERN Neutrino Experiment: Conclusions," ibid., pp. 584-590, the authors wrote, "Thus the ratio of neutral current elastic events is less than about 3%. Clearly neutral lepton currents cannot be admitted on a symmetrical basis with the charged." See also, M. M. Block, et al., Physics Letters, 12 (1964), pp. 281-285, on p. 285, for the same upper bound on neutral currents by these authors. Their result was also presented by D. C. Cundy, "Progress Report on Experimental Study of Neutrino Interactions in the CERN Heavy Liquid Bubble Chamber," in XII International Conference on High Energy Physics, Dubna, 5-15 August 1964 (Moscow: Atomizdat, 1966), pp. 7-15.

[21] Author's taped interview with P. Musset, 26 November 1980.

[22] D. C. Cundy, et al., "Upper Limits for Diagonal and Neutral Current Couplings in the CERN Neutrino Experiments," Physics Letters, 31B (1970), pp. 478-480.

[23] Letter: Lagarrigue to Jentschke, 12 April 1972.

[24] D. C. Cundy and C. Baltay, "Purely Leptonic Neutral Currents," CERN Technical Memorandum, 11 July 1972. Hereinafter, Technical Memorandum will be abbreviated TM.

[25]D. H. Perkins, "Neutral Currents," CERN TM to Gargamelle neutrino collaboration, undated, but before 28 April 1972 (see following note).

[26]Letter: Perkins to Baltay, 28 April 1972.

[27]Ibid.

[28]A. Pullia, in Neutrino '72, Balatonfured 11-17 June 1972, Volume 1 (Hungary: OMKDK-Technioinform, 1972), p. 229.

[29]C. Baltay, et al., "Proposal for a Meeting in Paris on Neutral Currents," CERN TM, 14 July 1972.

[30]C. Baltay, et al., "Work on Neutral Current Search at Milano and CERN," CERN TM, 14 July 1972.

[31]Ibid., pp. 3ff.

[32]D. H. Perkins, XVI International Conference on High Energy Physics, pp. 208-209. The reactor experiment referred to is the one reported on, for example, in H. S. Gurr, F. Reines, and H. W. Sobel, "Search for $\bar{\nu}_e + e^-$ Scattering," Physical Review Letters, 28 (1972), pp. 1406-1409. Antineutrinos interacted in a target of plastic scintillator and recoil electrons were detected. Severe background was encountered due to beta-decay.

[33]D. H. Perkins, XVI International Conference on High Energy Physics, pp. 208-209.

[34]Musset interview, Vialle interview. Also author's taped interview with D. C. Cundy, 27 November 1980.

[35]Lecture transparencies by P. Musset from APS meeting, January 1973.

[36]Musset interview.

[37]E. A. Paschos and L. Wolfenstein, "Tests for Neutral Currents in Neutrino Reactions," Physical Review D, 7 (1973), pp. 91-95.

[38]A. Rousset, "Schedule of Neutrino Experiments in Gargamelle in 1973," CERN TM, 19 February 1973.

[39]Ibid.

[40]P. Musset, "Study of Hadronic Neutral Currents," CERN TM, 19 March 1973.

[41]Ibid. Cf. P. Musset, addendum to report, CERN TM, 26 March 1973.

[42]D. C. Cundy, "Minutes of the Neutrino Collaboration Meeting on 21st March 1973 at CERN," CERN TM, 26 March 1973.

[43]J. P. Vialle and D. Blum, "Principe de Simulation du bruit de fond de neutrons dans Gargamelle," CERN TM, 15 May 1973. This work summarized the earlier computer programs.

[44]For example, in A. Rousset, "Calcul du bruit de fond de neutrons," CERN TM, 22 May 1973.

[45]Cundy interview.

[46]See original argument in A. Rousset, ibid., and the schematic one on which some of the above is based in A. Rousset, "Neutral Currents," Philadelphia, 1974 (full reference cited above in note 1).

[47]Vialle interview.

[48]Letter: Pullia to Musset, 11 July 1973.

[49]Cundy, for instance, did not feel the paper was sufficiently convincing at that time. Cundy interview, Musset interview, Vialle interview.

[50]F. J. Hasert, et al., "Observation of Neutrino-Like Interactions in the Gargamelle Neutrino Experiment," Physics Letters, 46B (1973), pp. 138-140. Received 25 July 1973.

[51]F. J. Hasert, et al., "Observation of Neutrino-Like Interactions Without Muon or Electron in the Gargamelle Neutrino Experiment," Nuclear Physics B, 73 (1974), pp. 1-22.

[52]A. K. Mann, "W Searches with High-Energy Neutrinos and High-Z Detectors," 1969 Summer Study, Volume 1. Particle Beams, Arthur Roberts, ed. (Batavia: National Accelerator Laboratory, 1969), pp. 201-207.

[53]Author's taped interview with A. K. Mann, 29 September 1980.

[54]U. Camerini, D. Cline, W. F. Fry, and W. M. Powell, "Search for Neutral Leptonic Currents in K^+ Decay," Physical Review Letters, 13 (1964), pp. 318-321.

[55]Ibid., p. 319.

[56]D. Cline, "Experimental Search for Weak Neutral Currents," reprinted from the school held in Heceg Novi, Yugoslavia, 1967.

[57]Ibid., pp. 27-28.

[58]D. Cline and A. K. Mann, "Preliminary Version of a Proposal for Neutrino Scattering Experiments at NAL," typed manuscript, 18 December 1969.

[59]Ibid.

[60]Mann interview.

[61]J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π Decay of the K_2^0 Meson," Physical Review Letters, 13 (1964), p. 138.

[62]Mann interview.

[63]E. W. Beier, D. Cline, A. K. Mann, J. Pilcher, D. D. Reeder, and C. Rubbia, NAL Neutrino Proposal, typescript, n.d. (but by the references and later documents which refer to it, the proposal must have been completed in early 1970).

[64]G. Danby, et al., Physical Review Letters, 9 (1962), pp. 36-44.

[65]D. Cline, A. K. Mann, and C. Rubbia, "Detection of the Weak Intermediate Boson Through Its Hadronic Decay Modes," Physical Review Letters, 25 (1970), pp. 1309-1312.

[66]S. Weinberg, "Unified Theory of Weak and Electromagnetic Interactions," Reviews of Modern Physics, 52 (1980), p. 518.

[67]Author's interview with S. Weinberg. Author's taped interview with C. Rubbia, 3 October 1980.

[68]Rubbia interview.

[69]Letter: A. Benvenuti, D. Cline, R. L. Imlay, W. Ford, A. K. Mann, J. E. Pilcher, D. D. Reeder, C. Rubbia, L. R. Sulak to R. R. Wilson, 14 March 1972.

[70]Rubbia interview.

[71]Mann interview.

[72]E. Beier, et al., "Search for Doubly Charged Weak Currents Through $K^+ \longrightarrow \pi^- e^+ \mu^+$," Physical Review Letters, 29 (1972), pp. 678-682.

[73]A. Benvenuti, et al., Physical Review Letters, 30 (1973), pp. 1084-1087.

[74]Author's taped interview with L. Sulak, 8 September 1980.

[75]This number is from the first version of "Observation of Muon-less Neutrino-Induced Inelastic Interactions," typescript, delivered by hand by L. Sulak to George Trigg (editor of Physical Review Letters) on 3 August 1973. A slightly revised version was submitted on 14 September 1973. The first paper will be referred to as "Early Version Observations," and the second one as LH 319, the number assigned it by Physical Review Letters when it was sent to the referees.

[76]Letter dated 17 July 1973.

[77]Letter of 18 July 1973 from Lagarrigue to Rubbia.

[78]Compare histograms on 400 GeV data on the early version of 3 August and a second version of 14 September.

[79]See Rubbia's comment after Myatt's talk at the Bonn Conference (cited below): "The important question in my opinion is whether neutral

currents exist or not, not so much the value of the branching ratio." Bonn, p. 405.

[80]Rubbia interview, Sulak interview. See also, D. Reeder, NAL TM, "Report of Bonn Conference": "The organizers had not scheduled a report by us nor would they schedule it in response to our request." The NAL group was permitted to present two short talks in the parallel sessions. Internal Report, 4 September 1973.

[81]D. Cline, "Performance of Revised E1A Detector for μless Event Search," NAL TM, 1 October 1973.

[82]Ibid.

[83]Author's taped interview with D. Cline, 14 January 1981.

[84]Mann interview.

[85]D. Cline, "Statistical Analysis of the μless Events from the Test Run," NAL TM, 11 October 1973.

[86]Ibid.

[87]Cline interview.

[88]D. Cline, "Unified Analysis of the μless Events in the November-December Runs," NAL TM, 16 October 1973.

[89]Anonymous Referee Reports sent with a cover letter from Physical Review Letters to L. Sulak, 16 October 1973.

[90]Letter from B. Aubert to Neutrino Collaboration at NAL, 16 October 1973.

[91]Typescript draft entitled, "Search for Neutrino Induced Events Without a Muon in the Final State." The manuscript is undated, but is referred to in a letter dated 13 November (discussed below) and was therefore probably written during the second week of November 1973.

[92]Ibid., p. 1.

[93]Mann interview.

[94]Letter: Musset to author, 9 April 1981; author's interview with A. Rousset, 30 November 1980.

[95]Rousset interview.

[96]P. Musset and J. P. Vialle, CERN TM, 20 November 1973.

[97]R. Imlay, "Calculation of Punch Through for the Horn Data," NAL TM, 29 November 1973.

[98]Letter from D. Cline, A. K. Mann, D. D. Reeder, and C. Rubbia to A. Lagarrigue, 13 November 1973, signed by Cline, Mann, and Rubbia. Signed version in Mann's files, unsigned copy found in Lagarrigue's scientific papers at Orsay. I would like to thank Mme. Lagarrigue and Professor Morrellet for permission to see these papers.

[99]L. Sulak, "Early Study of ν Run, November '73," NAL TM, 17 November 1973.

[100]Ibid.

[101]R. Imlay, "Calculation of Punch Through for the Horn Data," NAL TM, 29 November 1973.

[102] D. Cline, "Data Reported at NAL Talk, December 6, 1973," NAL TM, 13 December 1973. This TM includes photocopies of the transparencies shown at the talk.

[103] Ibid. See Figure 16.

[104] Ibid. See Figure 16.

[105] Mann interview.

[106] D. Cline, "Are We Seeing a μless Signal at the Level of 10%?" NAL TM, 13 December 1973.

[107] Ibid.

[108] Mann interview.

[109] A. K. Mann, "Summary of Rescan of 300 GeV $\bar{\nu}$ Run for N/C Experiment," NAL TM, 26 January 1974, p. 2.

[110] B. Aubert, F. Ling, and R. Imlay, no title, NAL TM, 18 December 1973.

[111] B. Aubert, et al., Physical Review Letters, 32 (1974), pp. 1454-1457.

CHAPTER V

AFTERWORD: HOW EXPERIMENTS END

To understand the decision to end an experiment we are forced to step outside of the division of scientific argument into a logic of discovery and a logic of justification. In Popper's scheme explicitly, and in other discussions implicitly, it is assumed that accounts of experiments must either be of the reconstructed type found in textbooks and review articles or else amount only to a description of the changing subjective beliefs of the experimentalists. In order to sketch a more accurate picture of the history of experimentation, we must insert between these two extremes a degree of experimental certainty one might call "plausible" or "reasonable." Plausible evidence could be so weak as merely to suggest that an experimentalist look somewhat longer at an effect, or be so strong as to lead to a radical alteration or abandonment of the experiment.

The three studies here illustrate some of the ways that evidence becomes plausible to the experimentalists in the course of their work, and how this evidence is eventually transformed into a form appropriate for the published literature. Einstein and de Haas anticipated their (incorrect) result and therefore accepted it as plausible far more readily than they might otherwise have done. Several factors seem to have played a role in shaping their theoretical expectations: Einstein's views on the Lorentz electron, his long-standing interest in the origin of the zero-point energy, and above all his sympathy with Ampère's hypothesis unifying permanent magnetism and electromagnetism. Similar considerations of unity and symmetry had guided his earlier work on relativity and light quanta.

Expectation was translated into confirmation in two steps.

First, Einstein and de Haas confirmed their theoretical prediction in a striking qualitative fashion: magnetization unambiguously induced the rotation of a suspended iron bar. The two physicists then faced the difficult task of eliminating the many systematic errors attendant upon such a delicate measurement. Thus, at a later stage of the experiment, systematic errors coincidentally combined to give a quantitative confirmation of the value of g that they expected to find. Because Einstein and de Haas expected g to equal unity, and because they had already seen that the qualitative effect confirmed their expectations, this quantitative confirmation seemed very compelling to them. They therefore ended their experiment and published the result.

Over the course of several decades, Barnett published a variety of values for g. In his early experiments of 1909 he found g to be many times higher than 2. In 1914, he determined that g was approximately 2. Soon afterwards, in 1917, he wrote that his latest experiments indicated that g was closer to 1. By the 1920's Barnett again placed the value of g as near to 2. As in the case of the Einstein and de Haas work, Barnett's mistaken results were closely tied to his theoretical predispositions. In his early work (before 1915), Barnett's principal interest was to demonstrate the qualitative fact that rotation would magnetize matter. This, he hoped, would provide a cause for terrestrial magnetism and explain the near coincidence of the geographic and magnetic North Poles. At the time Barnett had no particular quantitative prediction and soon confirmed the qualitative effect: rotation induced magnetization. Although his result was many times higher than what in retrospect it "should" have been, he had no compelling reason in 1909 to delay concluding those experiments. Only after Einstein's result that g = 1 was published in 1915 did Barnett's experiments begin to yield a similar result. Again, systematic errors, combined with a quantitative prediction, facilitated the decision to end the experiment when it confirmed the theory.

Barnett's decision to stop at that point probably was due more to the weight he attached to Einstein's work than to a considered reassessment of his theoretical priorities. I do not mean to imply that any of these experimentalists adjusted their answers to find a given result, but instead that theoretical and professional expectations necessarily play a role <u>both</u> when the experimentalist designs the experiment <u>and</u> when he decides to end the project.

Similarly, the developments leading up to the discovery of the muon attract our attention in part through the convergence of a variety of errors. Why was it that so many of the west coast American cosmic ray physicists had arrived at such apparently odd conclusions about the bands of cosmic ray energies, the absence of a latitude effect, the ionization maximum in the atmosphere, and so on? How did these experimentalists decide that their results were real and that their experiments were over? For Millikan the banded energies were in accord with his account of the continuous creation of the elements that would avoid the eventual heat-death of the universe. We have seen how the various incorrect experimental results exhibit a coherence; each was implied by Millikan's ideas about the origin of matter in interstellar space. Furthermore, like the oil-drop experiment (to which he explicitly likened his energy-band experiments) Millikan thought the electron spectrum, like the electric charge, would consist only of discrete values. In this respect his prior experimental experience in his protracted dispute with Ehrenhaft over the quantization of electric charge coincided with his general theoretical and philosophical convictions. This may account for the particular tenacity with which Millikan defended his cosmic ray theories.

A different kind of expectation may have been involved in Millikan's students' and associates' experimental confirmation of Millikan's ideas--they probably reflect more the force of Millikan's personality and reputation than shared beliefs. Perhaps they played roles similar to that of Barnett and de Haas when they confirmed Einstein's erroneous predictions.

In Millikan's group quantum mechanics played only a very peripheral role in the planning and interpretation of experiments. But elsewhere, by the 1930's, quantum mechanics was being used in an essential way to advance other domains of physics. Bethe, Rossi, Furry, Street, and many others were greatly impressed by the progress quantum mechanics afforded. Building on Bothe and Kolhörster's first experiments, Rossi designed experimental apparatus on the supposition that cosmic rays were charged corpuscles. This assumption motivated his development of logic circuits and the use of the circuits helped shape the kinds of experiments he undertook and the interpretation he eventually placed on his results. For example, when Rossi found a high rate of coincidence between two counters separated by large quantities of lead, he published the result as a demonstration that charged particles were extremely penetrating. By contrast Anderson, Millikan, Neddermeyer, and Pickering had a theoretical motivation for denying that charged particles could penetrate so far through lead. In addition, they had seen the striking increase in the size of showers as one increased the thickness of thin (<1 1/2 cm) lead plates. They ended the experiment with the assertion that Rossi had simply observed extended showers. This was understandable, considering they had what seemed at the time to be striking qualitative confirmation of a well-entrenched theoretical predisposition.

Three years later, Anderson and Neddermeyer and Street and Stevenson discovered a particle, intermediate in mass between the electron and the proton, that easily penetrated great quantities of lead. The two groups approached the problem from two very different theoretical and experimental frameworks. Anderson's work grew out of Millikan's program; Street's grew out of the work of Rossi and indirectly out of the quantum mechanical considerations of Bethe, Furry, and Oppenheimer. Consequently, the two groups employed quite different types of apparatus and were persuaded by different types of evidence that a new particle existed.

In the study of the more recent experiments of E1A and Gargamelle, we see how in addition to theoretical expectations the two styles of the statistical and exemplary forms of argumentation joined with the theoretical expectations in the decision to end the experiment. Some of the experimentalists in E1A had earlier conducted experiments based on statistical analyses of data. Others had performed bubble chamber experiments in the search for very rare events, where individual photographs of events were the object of investigation. In large part, Cline's doubts about the Monte-Carlo approach of the first E1A experiment must be seen as a continuation of an experimental style developed over years of successful experience with bubble chamber techniques. This would explain, at least in part, Cline's reluctance to end the experiment after the first paper was prepared and his satisfaction with the evidence when he found several events he could consider "golden."

More generally, it would be interesting to explore the historical roots of these two types of evidential arguments, the one based on statistics, the other on "golden events." Undoubtedly the use of statistics in physics experiments is tied to the evolution of the treatment of error in experiments. Together, statistics and errors stand centrally among a cluster of concerns surrounding the historical definition of a convincing experimental demonstration.

No reference to golden events was ever published by E1A as part of their demonstration of the existence of neutral currents, and for good reason. Evidence undergoes a transformation from the kind of private argument that persuades someone familiar with a particular set of techniques and apparatus, to the kind of evidence that is presented in published journal articles. Certainly this change occurred in experiments long before particle accelerators. But one inevitable consequence of the increase in size and duration of high energy experiments is that the division of labor forces each of several subgroups within the collaboration to find its own way to a conclusion. Of course information is shared among the different groups but each subgroup is necessarily most familiar with their own data and techniques. For Cline and for Mann these golden events were persuasive, but only later, with the termination of the experimental program that made use of more statistical analyses, was the group ready to publish their conclusions. It is a full time job to work as Vialle and Blum did on their Monte Carlo. They could not have been as familiar with the other Monte Carlo programs as Fry and Haidt, nor could they be as fluent with the thermodynamic equilibrium arguments as Rousset, nor as involved in the specific background analyses as Musset, Fiorini, or Pullia.

In a sense, the high energy physics experiment has become a meta-experiment standing above the many smaller subgroups performing tasks as intricate and extensive as entire experiments from earlier times. Indeed, one interesting feature of these subgroups is that their goals are often not identical to the overall question posed by the group--they are not miniatures of the whole. For instance the final goal of ElA was the investigation of the existence of weak neutral currents, whereas the subgroups had specific technical objectives such as determining punch-through. Indeed, many of the subgroups had as their main task the achievement of a better understanding of the apparatus itself.

In any experiment, recent or not, the apparatus must be understood before any conclusions can be drawn from the information it provides. Before the cosmic ray physicists were confident that turbulence in the cloud chamber gas was not a significant problem, the ionization tracks provided only equivocal evidence for the properties of the cosmic ray particles. Similarly, in the high energy experiments, as long as the detection efficiency in E1A was unknown, the observed muonless events could not be taken as strong evidence for neutral currents. Only as the participating physicists developed confidence in their understanding of the detailed behavior of the apparatus did the machines become a tool for the exploration of new physics.

Here too the difference between high energy and smaller scale experiments becomes apparent. Anderson, Street, or Rossi could achieve a "feel" for the cloud chambers and counter circuits they used. After working with certain kinds of equipment for many years, sources of error, inefficiencies, etc. become part of an almost intuitive or artisanal knowledge about an experiment. The same degree of certainty is manifestly not possible for any single individual in a large accelerator experiment. A subgroup working on the design and operation of the initial data reduction in the computer room may have little feeling for possible errors in the behavior of sparking inefficiencies in the detector. One observation heard over and again from young high energy experimental physicists is that their senior investigators have lost touch with the problems that emerge in the day-to-day operation of particular sections of the experiment. Those drawing conclusions about new particles from the reduced data are sometimes not aware of very specific machine characteristics that figure crucially in assessing the degree of certainty that ought to be ascribed to the conclusions. How much of a problem this may be depends on the details of individual experiments, but any account of how an experiment ends must take into consideration who is likely to be familiar with what in such a vast undertaking.

Theory, like experiment has greatly increased in complexity. One consequence of this is that the connection between the fundamental theory and the predictions for specific experimental arrangements have become much more elaborate. In the Einstein-de Haas experiments the theoretical calculations needed to pass from the theory that established $g = 2m/e$ to an experimental prediction for the resonant frequency of the iron rod were straightforward. They involved no more than a modified solution to the damped harmonic oscillator equation. Of course other experiments at the time involved more complicated applications of classical physics. Still, one of the interesting features of the cloud chamber experiments leading up to the discovery of the muon was the fundamental role played by calculations linking experiment to theory. Before Carlson and Oppenheimer's paper on shower phenomena, Anderson, Street, Rossi and many others treated the shower particles as problematic, and the penetrating particles as electrons. By quantitatively linking the showers with the Bethe-Heitler theory, Carlson and Oppenheimer effectively changed the problem to one of understanding the penetrating particles' properties. Without this theoretical work <u>between</u> the fundamental theory and the experimental phenomena, the experimental evidence left open the possibility that quantum electrodynamics was not valid at high energies.

In the weak neutral current case, an equally important role must be ascribed to the intermediate calculations. Without the calculations of Pais, Treiman, Paschos, and Wolfenstein, the hadronic muonless events could never have served as a confirmation or refutation of Weinberg-Salam theory. These calculations, in turn, rested upon the parton model. It will be important in any more comprehensive account of the evolution of the conduct of experiments in the twentieth century to trace the changing and increasingly subtle role of these intermediate calculations as the gap widens between experimental results and fundamental equations.

I would like to end these observations by suggesting some directions further research on the history of experimentation might take rather than by listing a set of over-arching generalizations about how experiments proceed. Many concerns intersect in the decision to end an experiment. Some of these are general and philosophical, some are specifically theoretical, and others are purely technical or practical. Therefore one line of inquiry leading back from the decision to end an experiment must include the theoretical expectations of the investigators. These may involve commitments to specific theories, aesthetic criteria, or philosophical beliefs. One could also identify individual or group preferences for certain types of evidence or styles of experimentation. We can also trace the experimentalists' changing familiarity and confidence in the detailed technology of the apparatus. Finally, there are factors altogether extrinsic to the scientific content of the experiment that nonetheless play a crucial role in the decision to end an experiment. These include the organization of the collaboration, the presence of competing groups, and constraints on when and how data can be collected.

We are still a long way from an understanding of how experiments have evolved through modern physics, but I hope these cases have demonstrated the necessity of asking the question, "How do experiments end?" as we write this history.

BIBLIOGRAPHY, CHAPTER II

Published Sources

Ampère, . "Mémoire sur la théorie mathématique des phénomènes électrodynamique uniquement déduite de l'expérience, dans lequel se trouvent réunis les Mémoires que M. Ampère a communiqués à l'Académie royale des sciences, dans les séances des 4 et 26 décembre 1820, 10 juin 1822, 22 décembre 1823, 12 septembre et 21 novembre 1825." Mémoires de l'académie royale des sciences de l'institut de France, VI (1823), issued 1827, pp. 175-388.

________. "Mémoire Présenté à l'Académie des Sciences, le 2 octobre 1820, où se trouve compris le resumé de ce qui avait été lu à la même Académie les 18 et 25 septembre 1820, sur les effets des courans électriques." Annales de Chimie et de Physique, 15 (1820), pp. 59-76 and pp. 170-218.

Arvidsson, G. "Eine Untersuchung über die Ampèreschen Molekularströme nach der Methode von A. Einstein und W. J. de Haas." Physikalische Zeitschrift, 21 (1920), pp. 88-91.

Barnett, S. J. "Magnetization By Angular Acceleration." Science, 30 (1909), p. 413. Dated 5 August 1909.

________. "Magnetization By Rotation." Physical Review, 6 (1915), pp. 239-270.

________. "Magnetization By Rotation." Physical Review, 6 (1915), pp. 171-172. Dated June 1915.

________. "The Magnetization of Iron, Nickel and Cobalt by Rotation and the Nature of the Magnetic Molecule." Physical Review, 10 (1917), pp. 7-21. Presented to APS December 1916.

________. "Further Experiments on Magnetization By Rotation." Proceedings Philosophical Society Washington, Journal Washington Academy of Sciences, 11 (1921), pp. 162-163. Presented 9 October 1920.

________. "The Angular Momentum of the Elementary Magnet." In "Theories of Magnetism." Bulletin of the National Research Council. Volume 3, part 3, number 18 (1922), pp. 235-268.

________. "Gyromagnetic Effects: History, Theory and Experiment." Physica, 13 (1933), pp. 241-268. Dated June 1933.

________, and Barnett, L. J. H. "Additional Experiments on the Nature of the Magnetic Molcule." Physical Review, 17 (1921), pp. 404-405.

________, and ________. "Improved Experiments on Magnetization by Rotation." Physical Review, 20 (1922), pp. 90-91. Abstract 17.

________, and ________. "New Researches on the Magnetization of Ferromagnetic Substances by Rotation and the Nature of the Elementary Magnet." Proc. of the American Academy of Arts and Sciences, 60 (1925), pp. 128-216. Received 14 May 1924.

Beck, E. "Zum experimentellen Nachweis der Ampèreschen Molekularströme." Annalen der Physik, 60 (1919), pp. 109-148. Received 7 May 1919.

Bohr, Niels. "On the Constitution of Atoms and Molecules." Phil. Mag., 26 (1913), pp. 1-25.

Chattock, A. P., and Bates, L. F. "On the Richardson Gyromagnetic Effect." Phil. Trans., 223 (1922), pp.

Compton, A. H. "The Magnetic Electron." Journal of the Franklin Institute, 192 (1921), pp. 145-155.

Ehrenfest, P. "Ersetzung der Hypothese vom unmechanischen Zwang durch eine Forderung bezüglich des inneren Verhaltens jedes einzelnen Elektrons." Die Naturwissenschaften, 47 (1925), pp. 953-954.

Einstein, A. "Die Plancksche Theorie der Strahlung und die theorie der spezifischen wärme." Annalen der Physik, 22 (1907)

________. "Über die Entwicklung unserer Anschauung uber das Wesen und die konstitution der Strahlung." Physikalische Zeitschrift, 10 (1909), pp. 817-826.

________. "Berichtigung zu meiner gemeinsamen mit Herrn W. J. de Haas veröffentlichten Arbeit 'Experimenteller Nachweis der Ampèreschen Molekularströme." Berichte der Detuschen Physikalischen Gesellschaft (1915), p. 203.

________. "Ein einfaches Experiment zum Nachweis der Ampèreschen Molekularströme." Verh. der Deutschen Physikalischen Gesellschaft, (Berichte 14), 18 (1916), pp. 173-177. Presented 25 February 1915.

________. Albert Einstein--Michele Besso Correspondence 1903-1955. Translated and edited by P. Speziali. Paris: Hermann, 1972.

________, and de Haas, W. J. "Experimenteller Nachweis der Ampèreschen Molekularströme." Verh. der Deutschen Physikalischen Gesellschaft, (Berichte 13), 17 (1915), pp. 152-170.

________, and ________. "Experimental Proof of the Existence of Ampère's Molecular Currents." Koninklijke Akademie van Wetenschappen Te Amsterdam, Proceedings Royal Academy of Amsterdam, 18 (1916), pp. 696-711. Communicated 23 April 1915.

________, and Hopf, L. "Statistische Untersuchung der Bewegung eines Resonators in einem Strahlungsfeld." Annalen der Physik, 32 (1910), pp. 1105-1115. Received 29 August 1910.

________, and Stern, O. "Einige Argumente für die Annahme einer molekularen Agitation beim absoluten Nullpunkt." Annalen der Physik, 40 (1913), pp. 551-560. Received 5 January 1913.

de Haas, W. J. "Further Experiments on the Moment of Momentum Existing in a Magnet." Koninklijke Akademie van Wetenschappen Te Amsterdam, 18 (1916), pp. 1281-1299. Communicated 25 September 1915.

________. "Weitere Versuche über die Realität der Ampèreschen Molekularströme." Verh. der Deutschen Physikalischen Gesellschaft, 18 (1916), pp. 423-443. Received 27 November 1916.

________. "Le Moment de la quantité de mouvement dans un corps magnetique." Atomes et Electrons. Institut Internationale de Physique Solvay. Rapports et discussions tenu à Bruxelles du 1er au 6 avril 1921. (Paris: Gauthier-Villars et C[ie], 1923), pp. 206-227.

________, and de Haas-Lorentz, G. L. "Een Proef van Maxwell en de Moleculaire Stroomen van Ampère." Amsterdam Koninklijke Akademie Verslag Wissen Naturkuunde, 24, 1 (1915), pp. 398-404.

Jenkins, F., ed. Reports of the Committee on Electrical Standards appointed by the British Association for the Advancement of Science. London: E. and F. N. Spon, 1873.

Maxwell, James Clerk. A Treatise on Electricity and Magnetism. Oxford: Clarendon Press, 1881.

Perry, J. Spinning Tops and Gyroscopic Motions. New York: Dover, 1957. Originally Spinning Tops (London, 1929).

Richardson, O. W. "A Mechanical Effect Accompanying Magnetization." Physical Review, 26 (1908), pp. 248-253.

________. The Electron Theory of Matter. Cambridge: Cambridge Univ. Press, 1914.

Schuster, A. "A Critical Examination of the Possible Causes of Terrestrial Magnetism." Proceedings of the Royal Society of London, 24 (1911-12), pp. 121-137.

Stewart, J. Q. "The Moment of Momentum Accompanying Magnetic Moment in Iron and Nickel." Physical Review, 11 (1918), pp. 100-120.

Sucksmith, C. N., and Bates, L. F. "On a Null Method of Measuring the Gyromagnetic Ratio." Proceedings of the Royal Society, 104 (1923), pp. 499-511. Received 24 July 1923.

Thomas, L. H. "The Motion of the Spinning Electron." Nature, 117 (1926), p. 514. Dated 20 February 1926.

Physics Reviews

Scott, G. G. "Theory of Gyromagnetic Effects and Some Related Phenomena." Reviews of Modern Physics, 34 (1962), pp. 143-165.

Stoner, E. C. Magnetism and Matter. London: Methuen and Co., Ltd., 1934.

Uhlenbeck, G. E., and Goudsmit, S. "Spinning Electrons and the Structure of Spectra." Nature, 117 (1926), pp. 264-265.

Historical Studies

Deutsch, M. "Evidence and Inference in Nuclear Research." Daedalus, (Fall 1958), pp. 88-98.

Einstein, A. "H. A. Lorentz, His Creative Genius and His Personality." In H. A. Lorentz Impressions of His Life and Work, ed., G. L. de Haas-Lorentz. Amsterdam: North-Holland, 1957.

Everitt, F. James Clerk Maxwell. New York: Charles Scribner's Sons, 1975.

Frenkel, V. Ia. "Kistorii effekta Einshteina-De Gaaza." Vspekhi fizicheskikh nauk, 128 (July 1979), pp. 545-557.

Hanson, N. R. Patterns of Discovery. Cambridge: Cambridge Univ. Press, 1975.

Holton, G. Thematic Origins of Scientific Thought. Cambridge: Harvard Univ. Press, 1973.

Kirsten, C., and Treder, Hans-Jurgen. Albert Einstein in Berlin 1913-1933. Volume I. Berlin: Akademie Verlag, 1979.

Klein, M. J. Paul Ehrenfest. Volume 1. The Making of a Physicist. Amsterdam: North-Holland, 1970.

Kuhn, T. S. "The Function of Measurement in Modern Physical Science." Isis, 168 (1961), pp. 161-193.

________. Black-Body Theory and the Quantum Discontinuity, 1894-1912. Oxford: Oxford Univ. Press, 1978.

McCormmach, R. "H. A. Lorentz and the Electromagnetic View of Nature." Isis, 61 (1970), pp. 459-497.

Miller, A. I. "On Lorentz's Methodology." British Journal for the Philosophy of Science, 25 (1974), pp. 33ff.

Pais, A. "Einstein and the Quantum Theory." Reviews of Modern Physics, 51 (1979), pp.

Portugal, F. H. "The History of the Department of Terrestrial Magnetism of the Carnegie Institution of Washington." Unpublished typescript.

Prins, J. A. "Wander Johannes de Haas." DSB, 5, p. 610.

Schaffner, K. "The Lorentz Theory of Relativity." American Journal of Physics, 37 (1969), pp. 498-513.

Strutt, Robert John. Life of John William Strutt. Madison: Univ. of Wisconsin Press, 1968.

Tolman, R. C., and Stewart, T. D. "The Electromotive Force Produced by the Acceleration of Metals." Physical Review, 8 (1916), pp. 97-116. Dated 25 February 1916.

Treder, H.-J. "A. Einstein: 'Einfache Methode zum Nachweis der Ampèreschen Molekularströme'." Wissenschaft und Fortschritt, 2 (1979), p. 53.

Tricker, R. A. R. Early Electrodynamics: The First Law of Circulation. Oxford: Pergamon Press, 1965.

Weill-Brushwicg, Adrienne R. "Paul Langevin." DSB, VIII, pp. 8-14.

Whittaker, E. A History of the Theories of Aether and Electricity. New York: Humanities Press, 1973.

Woodruff, A. E. "Wilhelm Eduard Weber." DSB, 14, pp. 203-209.

BIBLIOGRAPHY, CHAPTER III

Published Primary Sources

Anderson, C. D. "Space-Distribution of x-ray Photoelectrons From the K and L Atomic Energy-Levels." Physical Review, 35 (1930), pp. 1139-1145. Received 2 April 1930.

________. "The Apparent Existence of Easily Deflectable Positives." Science, 76 (1932), pp. 238-239. Dated 1 September 1932.

________. "Cosmic-Ray Positive and Negative Electrons." Physical Review, 44 (1933), pp. 406-416. Received 19 June 1933.

________. "The Production and Properties of Positrons." In Nobel Lectures Physics 1922-1941. Amsterdam: Elsevier, 1965, pp. 365-376. Nobel Prize Lecture of 1936.

________; R. A. Millikan; Seth Neddermeyer; and William Pickering. "The Mechanism of Cosmic-Ray Counter Action." Physical Review, 45 (1934), pp. 352-363. Received 26 December 1933.

________, and Seth H. Neddermeyer. "Fundamental Processes in the Absorption of Cosmic-Ray Electrons and Photons." In International Conference on Physics, London 1934. Volume 1: Nuclear Physics. Cambridge: Cambridge University Press, 1935, pp. 171-187.

________, and ________. "Energy-Loss and the Production of Secondaries by Cosmic-Ray Electrons." Physical Review, 46 (1934), p. 325. Abstract 14.

________, and ________. "Cloud Chamber Observations of Cosmic-Rays at 4300 Meters Elevation and Near Sea-Level." Physical Review, 50 (1936), pp. 263-271. Received 9 June 1936.

de Benedetti, S. "Directional Measurements on the Cosmic Rays Near the Geomagnetic Equator." Physical Review, 45 (1934), pp. 212-216. Received 25 November 1933.

Bethe, H. "Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie." Annalen der Physik, 5 (1930), pp. 325-400. Received 3 April 1930.

________. "Bremsformel für Elektronen relativistischer Geschwindigkeit." Zeitschrift für Physik, 76 (1932), pp. 293-299. Received 4 May 1932.

________, and W. Heitler. "On the Stopping of Fast Particles and on the Creation of Positive Electrons." Proceedings of the Royal Society, A146 (1934), pp. 83-112. Received 27 February 1934.

Bhaba, H. J., and W. Heitler. "Passage of Fast Electrons through Matter." Nature, 138 (1936), p. 401. Received 29 July 1936.

Blackett, P. M. S., and G. P. S. Occhialini. "Some Photographs of the Tracks of Penetrating Radiation." Proceedings of the Royal Society of London, 139A (1933), pp. 699-720. Received 7 February 1933.

Bohr, N. "On the theory of the Decrease of Velocity of Moving Electrified Particles on Passing through Matter." Philosophical Magazine, 25 (1913), pp. 10-31. Received August 1912.

Born, M. "On the Quantum Theory of the Electromagnetic Field." Proceedings of the Royal Society, A143 (1934), pp. 410-437. Received 9 August 1933.

Bothe, W., and W. Kolhörster. "Das Wesen der Höhenstrahlung." Zeitschrift für Physik, 56 (1929), pp. 751-777. Received 18 June 1929.

Bowen, I. S.; R. A. Millikan; and H. V. Neher. "A Very High Altitude Survey of the Effect of Latitude Upon Cosmic Ray Intensities and an Attempt at a General Interpretation of Cosmic-Ray Phenomena." International Conference on Physics, London, 1934. Volume 1. Cambridge, Cambridge University Press, 1935, pp. 206-224.

Breit, G., and John A. Wheeler. "Collision of Two Light Quanta." Physical Review, 46 (1934), pp. 1087-1091. Received 23 October 1934.

Carlson, J. F., and J. R. Oppenheimer. "On Multiplicative Showers." Physical Review, 51 (1936), pp. 220-231. Received 8 December 1936.

Compton, A. H. "Some Evidence Regarding the Nature of Cosmic Rays." Physical Review, 43 (1933), p. 382.

________. "A Geographic Study of Cosmic Rays." Physical Review, 43 (1933), pp. 387-403. Received 30 January 1933.

________, and H. A. Bethe. "Composition of Cosmic Rays." Nature, 134 (1934), pp. 734-735. Received 29 October 1934.

Furry, W. H., and J. F. Carlson. "Production of High Energy Electron Pairs." Physical Review, 44 (1933), pp. 237-238. Dated 1 July 1933.

Fussell, L. "Production and Absorption of Cosmic-Ray Showers." Bulletin of the American Physical Society, 12 (1937), pp. 13-14. Abstract 41.

Heitler, W. "Über die bei sehr schnellen Stössen emittierte Strahlung." Zeitschrift für Physik, 84 (1933), pp. 145-167. Received 4 June 1933.

________, and F. Sauter. "Stopping of Fast Particles with Emission of Radiation and the Birth of Positive Electrons." Nature, 132 (1933), p. 892.

Hoerlin, H. "Latitude Effect of Cosmic Radiation." Nature, 132 (1933), p. 61. Received 8 June 1933.

Jauncey, G. E. M. "Possible Origin of the x-particle." Physical Review, 52 (1937), p. 1256. Dated 1 December 1937.

Johnson, T. H. "Evidence that Protons are the Primary Particles of the Hard Component." Reviews of Modern Physics, 11 (1939), pp. 208-210.

McLennen, J. C., and E. F. Burton. "Some Experiments on the Electrical Conductivity of Air." Physical Review, 16 (1903), pp. 184-192. Abstract of paper presented 31 December 1902 to American Physical Society.

Millikan, R. A. Science and Life. Boston, The Pilgrim Press, 1924.

________. Evolution in Science and Religion. New York, Kennikat Press, 1973. Reprinted from Original. New Haven, Yale University Press, 1927.

________. Science and the New Civilization. New York, Charles Scribner's Sons, 1930.

________. "Present Status of Theory and Experiment as to Atomic Disintegration and Atomic Synthesis." Smithsonian Institution Annual Report 1930, pp. 277-285. Reprinted from Nature, 127 (1931). Address delivered 31 January 1931.

________. "Sur les Rayons Cosmiques." Annales de l'Institut Henri Poincaré, 3 (1932), pp. 447-464.

________. "New Techniques in the Cosmic-Ray Field and Some of the Results Obtained with Them." Physical Review, 43 (1933), pp. 661-669.

________. Electrons (+ and -), Protons, Photons, Neutrons, and Cosmic Rays. Chicago, University of Chicago Press, 1935.

________. Cosmic Rays. New York, Macmillan, 1939.

________. "The Present Status of the Evidence for the Atom-Annihilation." Reviews of Modern Physics, 21 (1949), pp. 1-13.

________, and C. D. Anderson. "Cosmic-Ray Energies and Their Bearing on the Photon and Neutron Hypotheses." Physical Review, 40 (1932), pp. 325-328. Received 12 April 1932.

________, and ________. "Cosmic-Ray Energies and Their Bearing on the Nature of These Rays." Physical Review, 40 (1932), p. 1056. Abstract 118.

________, and I. S. Bowen. "High Frequency Rays of Cosmic Origin I. Sounding Balloon Observations at Extreme Altitudes." Physical Review, 27 (1926), pp. 353-361. Received 24 December 1925.

________, and ________. "The Significance of Recent Cosmic-Ray Experiments." Proceedings of the National Academy of Sciences, 16 (1930), pp. 421-425. Read 29 April 1930.

________, and G. H. Cameron. "High Frequency Rays of Cosmic Origin III. Measurements in Snow-Fed Lakes at High Altitudes." Physical Review, 28 (1926), pp. 851-868. Received 7 August 1926.

________, and ________. "Direct Evidence of Atom-Building." Science, 67 (1928), p. 402.

________, and ________. "New Precision in Cosmic Ray Measurements; Yielding Extension of Spectrum and Indications of Bands." Physical Review, 31 (1928), pp. 921-930. Received 19 March 1928.

________, and ________. "The Origin of the Cosmic Rays." Physical Review, 32 (1928), pp. 533-557. Received 9 July 1928.

________, and ________. "Evidence that the Cosmic Rays Originate in Interstellar Space." Proceedings of the National Academy of Sciences, 14 (1928), pp. 637-650. Communicated 12 July 1928.

________, and R. M. Otis. "High Frequency Rays of Cosmic Origin II. Mountain Peak and Airplane Observations." Physical Review, 27 (1926), pp. 645-658. Received 1 March 1926.

Montgomery, C. G., and D. D. Montgomery. "The Heavy Particle Component of the Cosmic Radiation." Physical Review, 50 (1936), pp. 975-976. Dated 26 October 1936.

Neddermeyer, S. H. "The Penetrating Cosmic-Ray Particles." Physical Review, 53 (1938), pp. 102-103. Dated 14 December 1937.

________, and C. D. Anderson. "Note on the Nature of Cosmic-Ray Particles." Physical Review, 51 (1937), pp. 884-886. Received 30 March 1937.

________, and ________. "III. Composition of Cosmic Rays, Nature of Cosmic-Ray Particles." Reviews of Modern Physics, 11 (1939), pp. 191-210.

Nordheim, L. W. "Probability of Radiative Processes for Very High Energies." Physical Review, 49 (1936), pp. 189-191. Received 3 December 1935.

________. "The Absorption of Cosmic Rays in the Atmosphere." Bulletin of the American Physical Society, 12 (1937), p. 14. Abstract 42.

Oppenheimer, J. R. "Are the Formulae for the Absorption of High Energy Radiations Valid?" Physical Review, 47 (1935), pp. 44-52. Received 12 November 1934.

________, and M. S. Plesset. "On the Production of the Positive Electron." Physical Review, 45 (1934), pp. 53-55. Received 9 June 1933.

________, and R. Serber. "Note on the Nature of Cosmic-Ray Particles." Physical Review, 51 (1937), p. 1113. Dated 1 June 1937.

Rossi, Bruno. "Method of Registering Multiple Simultaneous Impulses of Several Geiger's Counters." Nature, 125 (1930), p. 636.

________. "On the Magnetic Deflection of Cosmic Rays." Physical Review, 36 (1930), p. 606. Received 3 July 1930.

________. "Über die Eigenschaften der durchdringenden Korpuskularstrahlung im Meeresniveau." Zeitschrift für Physik, 82 (1933), pp. 151-178. Received 24 February 1933.

________. "Some Results Arising from the Study of Cosmic Rays." International Conference on Physics. Cambridge, Cambridge Univ. Press, 1934, pp. 233-247.

Rutherford, E., and H. L. Cooke. "A Penetrating Radiation from the Earth's Surface." Physical Review, 16 (1903), p. 183. Abstract of paper presented 31 December 1902 to American Physical Society.

Stevenson, E. C., and J. C. Street. "Nature of the Penetrating Cosmic Radiation at Sea Level." Physical Review, 47 (1935), p. 643. Abstract 32.

________, and ________. "Cosmic-Ray Showers Produced by Electrons." Physical Review, 48 (1935), pp. 464-465. Dated 27 July 1935.

________, and ________. "Cloud Chamber Photographs of Counter Selected Cosmic-Ray Showers." Physical Review, 49 (1936), pp. 425-428. Received 24 January 1936.

________, and ________. "Cloud Chamber Photographs of Counter Selected Cosmic-Ray Showers." Physical Review, 49 (1936), p. 638. Abstract 8.

Stueckelberg, E. C. G. "On the Existence of Heavy Electrons." Physical Review, 52 (1937), pp. 41-42. Dated 6 June 1937.

Street, J. C. "Cloud Chamber Studies of Cosmic Ray Showers and Penetrating Particles." Journal of the Franklin Institute, 227 (1939), pp. 765-788.

________; E. G. Schneider; and E. C. Stevenson. "Heavy Particles from Lead." Physical Review, 48 (1935), p. 463. Dated 2 August 1935.

________, and E. C. Stevenson. "The Absorption of Cosmic-Ray Electrons." Physical Review, 47 (1935), pp. 891-895. Received 5 April 1935.

________, and ________. "Design and Operation of the Counter Controlled Cloud Chambers." Physical Review, 49 (1936), p. 638. Abstract 7.

________, and ________. "Design and Operation of Counter Controlled Cloud Chambers." Review of Scientific Instruments, 7 (1936), pp. 347-353.

________, and ________. "Penetrating Corpuscular Component of the Cosmic Radiation." Physical Review, 51 (1937), p. 1005. Abstract 40.

________, and ________. "Evidence for the Existence of a Particle of Mass Intermediate Between the Proton and Electron." Physical Review, 52 (1937), pp. 1003-1004. Dated 6 October 1937.

________, and R. H. Woodward. "Counter Calibration and Cosmic-Ray Intensity." Physical Review, 46 (1934), pp. 1029-1034. Received 1 October 1934.

________, and R. T. Young. "Transition Effects in the Cosmic Radiation." Physical Review, 46 (1934), pp. 823-824. Dated 27 September 1934.

________, and ________. "Shower Groups in Cosmic Radiation." Physical Review, 47 (1935), pp. 572-573. Dated 14 March 1935.

v. Weiszäcker. "Austrahlung bei Stössen sehr schneller Elektronen." Zeitschrift für Physik, 88 (1934), pp. 612-625. Received 28 February 1934.

Williams, E. J. "Applications of the Method of Impact Parameter." Proceedings of the Royal Society, 139A (1933), pp. 163-186. Received 7 November 1932.

________. "Nature of the High Energy Particles of Penetrating Radiation and Status of Ionization and Radiation Formulae." Physical Review, 45 (1934), pp. 729-730. Received 16 April 1934.

________, and E. Pickup. "Heavy Electrons in Cosmic Rays." Nature, 141 (1938), pp. 684-685. Dated 28 March 1938.

Woodward, R. H., and J. C. Street. "The Absorption of Cosmic-Ray Electrons in Lead." Physical Review, 47 (1935), p. 643. Abstract 31.

________, and ________. "Production and Absorption of Cosmic-Ray Showers." Physical Review, 47 (1935), p. 800. Abstract 75.

________, and ________. "The Absorption of Cosmic-Ray Electrons at 10,600 Ft. and at Sea Level." Physical Review, 49 (1936), p. 198. Abstract 22.

Young, R. T., and J. C. Street. "Cosmic-Ray Ionizations under Various Thicknesses of Lead Shield in Northern and Equatorial Latitudes at Different Altitudes." Physical Review, 51 (1937), p. 386. Abstract 64.

Secondary Sources

Ahlen, S. P. "Theoretical and Experimental Aspects of the Energy Loss of Relativistic Heavily Ionizing Particles." Reviews of Modern Physics, 52 (1980), pp. 121-173.

Bernstein, J. Hans Bethe, Prophet of Energy. New York, Basic Books, 1979.

Bjorken, J. D., and S. D. Drell. Relativistic Quantum Mechanics. New York, McGraw-Hill, 1964.

Bromberg, J. "The Impact of the Neutron: Bohr and Heisenberg." Historical Studies in the Physical Sciences, 3 (1971), pp. 307-341.

________. "The Concept of Particle Creation before and after Quantum Mechanics." Historical Studies in the Physical Sciences, 7 (1976), pp. 161-191.

Brown, L. "The Prediction and the Discovery of 'Yukawa's Meson.' Part 1. The Prediction." Centaurus, 25 (1981), pp. 71-132.

Hanson, N. R. The Concept of the Positron. Cambridge, Cambridge University Press, 1963.

Kargon, R. H. "The Conservative Mode: Robert A. Millikan and the Twentieth-Century Revolution in Physics." Isis, 68 (1977), pp. 509-526.

________. "Birth Cries of the Elements: Theory and Experiment along Millikan's Route to Cosmic Rays." In The Analytic Spirit, ed., H. Woolf. Ithaca, Cornell University Press, 1981.

Kevles, D. J. The Physicists. New York, Knopf, 1978.

________. "Robert A. Millikan." Scientific American, January 1979, pp. 142-151.

Seidel, R. W. Physics Research in California: The Rise of a Leading Sector in American Physics. Unpublished Ph.D. dissertation, University of California, Berkeley, 1978.

Smith, A., and C. Weiner. Robert Oppenheimer, Letters and Reflections. Cambridge: Harvard University Press, 1980.

Stuewer, R. The Compton Effect. New York, Science History Publications, 1975.

________, ed. Nuclear Physics in Retrospect, Proceedings of a Symposium on the 1930's. Minneapolis, University of Minnesota Press, 1979.

Retrospectives by Participants

Anderson, C. D. "Early Work in the Positron and Muon." American Journal of Physics, 29 (1961), pp. 825-830.

________. "Unraveling the Particle Content of the Cosmic Rays, Including Discovery of the Positron and Mu Meson." Paper for the International Symposium on Particle Physics, Fermilab, 1980.

Auger, P. What Are Cosmic Rays? Chicago, University of Chicago Press, 1944.

Bernardini, G. "The Discovery of μ-meson Among the Cosmic Rays." Talk delivered at International Symposium on Particle Physics, Fermilab, 1980.

Millikan, R. A. The Autobiography of Robert A. Millikan. New York, Prentice-Hall, 1950.

Rossi, B. Cosmic Rays. New York, McGraw-Hill, 1964.

BIBLIOGRAPHY, CHAPTER IV

Published Sources: Primary and Secondary

Abers, E. S., and Lee, B. W. "Gauge Theories." Physics Reports, 9 (1973), pp. 1-141. Received 5 April 1973.

Aiff-Steinberger, C., et al. "K_S and K_L Interference in the $\pi^+\pi^-$ Decay Mode, CP Invariance and the K_S-K_L Mass Difference." Physics Letters, 20 (1966), pp. 207-211. Received 28 December 1965.

Aubert, B., et al. "Further Observation of Muonless Neutrino-Induced Inelastic Interactions." Physical Review Letters, 32 (1974), pp. 1454-1457. Received 19 March 1974.

Baltay, C. "Neutrino Interactions II: Neutral Current Interactions and Charm Production." Proceedings of the 19th International Conference in High Energy Physics. Physical Society of Japan: Tokyo, 1978, pp. 882-903.

Beier, E., Buchholz, D. A., Mann, A. K., and Parker, S. H. "Search for Doubly Charged Weak Currents Through $K^+ \rightarrow \pi^- e^+ \mu^+$." Physical Review Letters, 29 (1972), pp. 678-682.

Bell, J. S., Løvseth, J., and Veltman, M. "CERN Neutrino Experiment: Conclusions." Proceedings of the Sienna International Conference on Elementary Particles 30 September-5 October 1963. Volume 1. Bologna: Società Italiana di Fisica, 1963, pp. 584-590.

Benvenuti, A., Cheng, D. C., Cline, D., Ford, W. T., Imlay, R., Ling, T. Y., Mann, A. K., Messing, F., Piccioni, R. L., Pilcher, J., Reeder, D. D., Rubbia, C., Stefanski, R., and Sulak, L. "Observation of Muonless Neutrino-Induced Inelastic Interactions." Physical Review Letters, 32 (1974), pp. 1084-1087. Received 3 August 1973. Appeared 8 April 1974.

Benvenuti, A., et al. "A Liquid-Scintillator Total-Absorption Hadron Calorimeter For the Study of Neutrino Interactions." Nuclear Instruments and Methods, 125 (1975), pp. 447-456. Received 5 February 1975.

Benvenuti, A., et al. "A Large Area Magnetic Spectrometer for the Study of High-Energy Neutrino Interactions." Nuclear Instruments and Methods, 125 (1975), pp. 457-460. Received 5 February 1975.

Bjorken, J. D., and Paschos, E. A. "Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon." Physical Review, 185 (1969), pp. 1975-1982. Received 10 April 1960.

Block, M. M., et al. "Neutrino Interactions in the CERN Heavy Liquid Bubble Chamber." Physics Letters, 12 (1964), pp. 281-285. Received 21 September 1964.

Briedenbach, M., et al. "Observed Behavior of Highly Inelastic Electron-Proton Scattering." Physical Review Letters, 23 (1969), pp. 935-939. Received 22 August 1969.

Camerini, V., Cline, D., Fry, W. F., and Powell, W. M. "Search For Neutral Leptonic Currents in K^+ Decay." Physical Review Letters, 13 (1964), pp. 318-321. Received 4 August 1964.

Camerini, V., Ljung, D., Sheaff, M., and Cline, D. "Experimental Search for Semileptonic Neutrino Neutral Currents." Physical Review Letters, 23 (1969), pp. 326-329. Received 20 June 1969.

Christenson, J. H., Cronin, J. W., Fitch, V. L., and Turlay, R. "Evidence for the 2 Decay of the K_2^0 Meson." Physical Review Letters, 13 (1964), pp. 138-140. Received 10 July 1964.

Cline, D. "Experimental Search for Weak Neutral Currents." Reprinted from the school held in Heceg Novi, Yugoslavia, 1967.

Cline, D. Preface to Unification of Elementary Forces and Gauge Theories. Papers presented at Ben Lee Memorial International Conference on Parity Nonconservation and Gauge Theories. Fermi National Accelerator Laboratory. October 20-22, 1977. London: Harwood Academic Publishers, 1978, pp. ix-xii.

Cline, D., and Fry, W. F. "Neutrino Scattering and New-Particle Production." Annual Review of Nuclear Science, 27 (1977), pp. 209-278.

Cline, D., Mann, A. K., and Rubbia, C. "Detection of the Weak Intermediate Boson Through Its Hadronic Decay Modes." Physical Review Letters, 25 (1970), pp. 1309-1312. Received 24 July 1970.

Cundy, D. C. Presentation of work of M. M. Block, et al., "Progress Report on Experimental Study of Neutrino Interactions in the CERN Heavy Liquid Bubble Chamber." In XII International Conference on High Energy Physics, Dubna, 5-15 August 1964. Moscow: Atomizdat, 1966, pp. 7-15.

Cundy, D. C., Myatt, G., Nezrick, F. A., Pattison, J. B. M., Perkins, D. H., Ramm, C. A., Venus, W., and Wachsmuth, H. W. "Upper Limits For Diagonal and Neutral Current Couplings in the CERN Neutrino Experiments." Physics Letters, 31b (1970), pp. 478-480. Received 16 February 1970.

Danby, G., Gaillard, J-M, Goulianos, K., Lederman, L. M., Mistry, N., Schwartz, M., and Steinberger, J. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos." Physical Review Letters, 9 (1962), pp. 36-44. Received 15 June 1962.

Feynman, R. High Energy Collisions. Third International Conference, State University of New York, Stony Brook, 1969. C. N. Yang, J. A. Cole, M. Good, R. Hwa, and J. Lee-Franzini, eds. New York: Gordon and Breach, 1969.

Franklin, A. "The Discovery and Nondiscovery of Parity Nonconservation." Studies in the History and Philosophy of Science, 10 (1979), pp. 201-257.

Glashow, S. L. "Towards a Unified Theory: Threads in a Tapestry." Reviews of Modern Physics, 52 (1980), pp. 539-543.

Glashow, S. L., and Georgi, H. "Unified Weak and Electromagnetic Interactions Without Neutral Currents." Physical Review Letters, 28 (1972), pp. 1494-1497. Received 13 March 1972.

Gurr, H. S., Reines, F., and Sobel, H. W. "Search for $\bar{\nu}_e + e^-$ Scattering." Physical Review Letters, 28 (1972), pp. 1406-1409.

Hasert, F. J., Faissner, H., Krenz, W., Von Krogh, J., Lanske, D., Morfin, J., Schultze, K., and Weertz, H. III Physikalisches Institut der technischen Hochschule, Aachen, Germany.

Bertrand-Coremans, G. H., Lemonne, J., Sacton, J., van Doninck, W., and Vilain, P. Interuniversity Institute for High Energies, U.I.B., V.U.B., Brussels, Belgium.

Brisson, V., Degrange, B., Haguenauer, H., Kluberg, L., Nguyen-Khac, U., and Petiau, P. Laboratoire de Physique des Hautes Energies, École Polytechnique, Paris, France.

Bellotti, E., Bonetti, S., Cavalli, D., Conta, C., Fiorini, E., and Rollier, M. Instituto di Fisica dell'Università, Milano and I.N.F.N., Milano, Italy.

Aubert, B., Chounet, L. M., Heusse, P., Lagarrigue, A., Lutz, A. M., and Vialle, J. P. Laboratoire de l'Accelerateur Linéaire, Orsay, France.

Bullock, F. W., Esten, M. J., Jones, T., McKenzie, J., Michette, A. G., Myatt, G., Pinfold, J., and Scott, W. J. University College of London, England.

"Search For Elastic Muon Neutrino Electron Scattering." Physics Letters, 46b (1973), pp. 121-124. Received 2 July 1973.

Hasert, F. J., et al. "Observation of Neutrino-Like Interactions in the Gargamelle Neutrino Experiment." Physics Letters, 46b (1973), pp. 138-140. Received 25 July 1973.

Hasert, F. J., et al. "Observation of Neutrino-Like Interactions Without Muon or Electron in the Gargamelle Neutrino Experiment." Nuclear Physics B, 73 (1974), pp. 1-22.

Kabir, P. K., ed. The Development of Weak Interaction Theory. New York: Gordon and Breach, 1963.

Kim, J. E., et al. "A Theoretical and Experimental Review of the Weak Neutral Current: A Determination of Its Structure and Limits on Deviations from the Minimal SU(2) x U(1) Electroweak Theory." Reviews of Modern Physics, 53 (1981), pp. 211-252.

Lee, B. W. "The Process $\nu_\mu + p \longrightarrow \nu_\mu + p + \pi^0$ in Weinberg's Model of Weak Interactions." Physics Letters, 40b (1972), pp. 420-422. Received 29 May 1972.

Lee, T. D., and Yang, C. N. "Theoretical Discussions on Possible High-Energy Neutrino Experiments." Physical Review Letters, 4 (1960), pp. 307-311. Received 23 February 1960.

Lee, T. D., and Yang, C. N. "Implications of the Intermediate Boson Basis of the Weak Interactions: Existence of a Quartet of Intermediate Bosons and Their Dual Isotopic Spin Transformation Properties." Physical Review, 119 (1960), pp. 1410-1419. Received 11 April 1960.

Lee, W. "Experimental Limit on the Neutral Current in the Semileptonic Processes." Physics Letters, 40b (1972), pp. 423-425. Received 29 May 1972.

Llewyn Smith, C. C. "Unified Models of Weak and Electromagnetic Interactions." Proceedings of the 6th International Symposium on Electron and Photon Interactions at High Energies. Amsterdam: North Holland Press, 1974, pp. 449-465.

Lutz, A. M. "Experimental Research of Neutral Currents in Interactions." Interactions faibles et électromagnétiques. Vol. 1: Huitieme Rencontre de Compte rendu de la Moriond. 4-16 Mars 1973.

Mann, A. K. "W Searches with High-Energy Neutrinos and High-Z Detectors." 1969 Summer Study. Volume 4: Particle Beams. Batavia: National Accelerator Laboratory, pp. 201-207.

Mann, A. K. "Present Status of Inelastic and Elastic Hadronic Weak Neutral Currents." In Unification of Elementary Forces and Gauge Theories. Papers presented at the Ben Lee Memorial International Conference on Parity Nonconservation and Gauge Theories. Fermi National Accelerator Laboratory, October 20-22, 1977. London: Harwood Academic Publishers, 1978, p. 20.

Musset, P. "Neutrino Interactions." Aix-en-Provence International Conference on Elementary Particles, 6 September 1973.

Musset, P. "État des résultats expérimentaux sur la recherche des courants neutres." Journal de Physique, 34 (1973), pp. C3-1 to C3-7. (Colloque C3 supplément au N° 11-12.)

Myatt, G. "Neutral Currents." In Proceedings of the 6th International Symposium on Electron and Photon Interactions at High Energies. Amsterdam: North-Holland Press, 1974, pp. 389-406.

Pais, A., and Treiman, S. B. "Neutral-Current Effects in a Class of Gauge Field Theories." Physical Review D, 6 (1972), pp. 2700-2703. Received 2 June 1972.

Paschos, E. A., and Wolfenstein, L. "Tests for Neutral Currents in Neutrino Interactions." Physical Review D, 7 (1973), pp. 91-95. Received 9 August 1972.

Perkins, D. H. Presenting work of H. H. Bingham, et al. "CERN Neutrino Experiment-Preliminary Bubble Chamber Results." Proceedings of the Sienna International Conference on Elementary Particles 30 September-5 October 1963. Volume 1. Bologna: Società Italiana di Fisica, 1963, pp. 555-584.

Perkins, D. H. "Neutrino Interactions." Proceedings of the XVI International Conference on High Energy Physics, 6 September-13 September 1972. Volume 4. Batavia: National Accelerator Laboratory, 1972, pp. 189-247.

Pullia, A. "Search for Neutral Currents in 'Gargamelle.'" [xt, July 1972, presented at Belaton, Hungary.] In Neutrino '72, Belatonfured 11-17 June 1972. Volume 1. Hungary: OMKDK-Technioinform, 1972, p. 229.

Reines, F., and Gurr, H. S. "Upper Limit for Elastic Scattering of Electron Antineutrinos by Electrons." Physical Review Letters, 24 (1970), pp. 1448-1452. Received 27 April 1970.

Rousset, A. "Neutral Currents." Talk presented at the International Conference on Neutrino Physics and Astrophysics, Philadelphia, 26-28 April 1974 and printed in Neutrino-1974, C. Baltay, ed., New York: American Institute of Physics, 1974, pp. 141-165.

Salam, A. "Weak and Electromagnetic Interactions." In Elementary Particle Theory, ed. N. Svartholm. Stockholm: Almquist and Forlag, 1968.

Salam, A. "Gauge Unification of Fundamental Forces." Reviews of Modern Physics, 52 (1980), pp. 525-536.

Schwartz, M. "Feasibility of Using High Energy Neutrinos to Study the Weak Interactions." Physical Review Letters, 4 (1960), pp. 306-307. Received 23 February 1960.

Sullivan, D., et al. "Understanding Rapid Theoretical Change in Particle Physics: A Month-By-Month Co-Citation Analysis." Scientometrics, 2 (1980), pp. 309-318.

t'Hooft, G. "Renormalization of Massless Yang-Mills Fields." Nuclear Physics B, 33 (1971), pp. 173-199. Received 12 February 1971.

t'Hooft, G. "Prediction for Neutrino-Electron Cross-Sections in Weinberg's Model of Weak Interactions." Physics Letters, 37b (1971), pp. 195-196. Received 27 October 1971.

Wachsmuth, H. "Latest Results from Gargamelle Neutrino Experiments." Proceedings of Summer Institute on Particle Physics at SLAC, SLAC 167, Vol. II, p. 235.

Wachsmuth, H. "Review of Latest Results from High-Energy Neutrino Experiments." Conference of the American Physical Society Division of Particles and Fields, Berkeley, 13-17 August 1973.

Weinberg, S. "A Model of Leptons." Physical Review Letters, 19 (1967), pp. 1264-1266. Received 17 October 1967.

Weinberg, S. "Effects of a Neutral Intermediate Boson in Semileptonic Processes." Physical Review D, 5 (1972), pp. 1412-1417. Received 6 December 1971.

Weinberg, S. "Recent Progress in Gauge Theories of the Weak, Electromagnetic and Strong Interactions." Reviews of Modern Physics, 46 (1974), pp. 255-277.

Weinberg, S. "Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions." Reviews of Modern Physics, 52 (1980), pp. 515-523.

Technical Memoranda from CERN (in Chronological Order)
(XT will abbreviate xeroxed typescript)

Lagarrigue, A., Musset, P., Rousset, A. "Projet de chambre à bulles à liquides lourds de 17m3" [CERN TM, XT, February 1964].

Allard, J. F., et al. "Proposition de construction d'une grande chambre à bulles à liquides lourds destinée à fonctionner auprès du synchrotron à protons du CERN" [CERN TM, XT (Summer 1964?)].

Perkins, D. H. "Draft Gargamelle Proposal" [CERN TM, XT, 4 February 1970].

Aubert, B., et al. "Amended Draft to the Gargamelle Proposal" [CERN TM, XT, 25 February 1970].

Aubert, B. to Perkins, D. H., 27 February 1970 [cover letter to "Amended Draft to the Gargamelle Proposal"].

Cundy, D. C. "Neutrino Collaboration-11th November 1971" [CERN TM, XT, 29 November 1971].

Cundy, D. C. "Resume of Neutrino Collaboration Meeting Held on 19th January 1972" [CERN TM, XT, 31 January 1972].

Perkins, D. H. "Neutral Currents" [CERN TM, XT, addressed to Gargamelle Collaboration, undated but before April 1972].

Lagarrigue, A. to Jentschke, W. K., 12 April 1972 [at least 3 copies made].

Perkins, D. H. to Baltay, C., 28 April 1972 [xeroxed copies to Gargamelle Collaboration].

Cundy, D. C., and Baltay, C. "Purely Leptonic Neutral Currents" [CERN TM, XT technical memorandum, 11 July 1972].

Baltay, C., Camerini, U., Fry, W., Musset, P., Osculati, B., and Pullia, A. "Memorandum to Collaborators Working on Neutral Currents, Subject: Proposal for a Paris Meeting on Neutral Currents" [CERN TM, XT, 14 July 1972].

Baltay, C., Camerini, U., Fry, W., Musset, P., Osculati, B., and Pullia, A. "Memorandum. Subject: Work on Neutral Current Search at Milano and CERN" [CERN TM, XT, 14 July 1972].

Aubert, B., Heusse, Ph., Kluberg, L., and Petiau, P. to Gargamelle Collaboration [CERN TM, XT, 17 July 1972].

Musset, P. [transparencies from presentation to American Physical Society Meeting, January 1972, in New York].

Cundy, D. C. "Neutrino Collaboration Meeting 30th January 1973" [CERN TM, XT, 6 February 1973].

Rousset, A. to Professor Cresti, "Memorandum. Schedule of Neutrino Experiments in Gargamelle in 1973" [CERN TM, XT, 19 February 1973].

Musset, P. "Study of Hadronic Neutral Currents" [CERN TM, XT, 19 March 1973].

Musset, P. to Neutrino Collaboration [CERN TM, XT, 26 March 1973].

Cundy, D. C. "Minutes of the Neutrino Collaboration Meeting Held on 21st March 1973 at CERN" [CERN TM, XT, 26 March 1973].

Musset, P. to Neutrino Collaboration [CERN TM, XT, 26 March 1973].

Musset, P. "List of the NC Events > 1 GeV Controlled at CERN on the 12 & 13 April 73" [CERN TM, XT, 17 April 1973].

Cundy, D. C., Musset, P., and Haidt, D. "Minutes of the Neutrino Collaboration Meeting Held at CERN on 11th & 12th April 73" [CERN TM, XT, 17 April 1973].

Vialle, J. P., and Blum, D. "Principe de simulation du bruit de fond de neutrons dans Gargamelle" [CERN TM, XT, 15 May 1973].

Rousset, A. "Calcul du bruit de fond de neutrons" [CERN TM, XT, 22 May 1973].

Fry, W., and Haidt, D. "Evaluation of the Neutron Flux in Equilibrium with the Neutrino Beam" [CERN TM, XT, 22 May 1973].

Aubert, B., Bullock, F. W., and Pullia, A. "Proposed Criteria for Linking Events on the Same Frame" [CERN TM, XT, 21 May 1973].

Musset, P. "Minutes of the Meeting of 16-17 May 1973 on the Checks of NC and CC" [CERN TM, XT, 21 May 1973].

Wachsmuth, H. "Radial $\nu(\bar{\nu})$ Flux and Event Distributions in 1971 ν and $\bar{\nu}$ Runs and NC/CC Ratios" [CERN TM, XT, 23 May 1973].

Wachsmuth, H. "Neutron Cascade Calculations in Iron and Freon Using Ranft's Program Tranka" [CERN TM, XT, 24 May 1973].

[author unknown]. "UCL Hadronic Neutral Current Data" [CERN TM, XT, 4 June 1973].

Fiorini, E. to Musset, P., 11 July 1973 [typed letter].

Blum, D. "Simulation dans Gargamelle d'un bruit de fond de neutrons produits par le faisceau de neutrons" [CERN TM, XT, 11 September 1973].

Musset, P. and Vialle, J. P. to Gargamelle Collaboration, 20 November 1973 [CERN TM, XT].

Rousset, A. to Weisskopf, V., 8 May 1974 [copies circulated to Gargamelle Collaboration].

Technical Memoranda from NAL (in Chronological Order)

Cline, D. and Mann, A. K. "Preliminary Version of a Proposal for Neutrino Scattering Experiments at NAL" [XT, 18 December 1969].

Beier, E. W., Cline, D., Mann, A. K., Pilcher, J., Reeder, D. D., and Rubbia, C. "Harvard-Pennsylvania-Wisconsin Collaboration NAL Neutrino Proposal" [proposal, XT, 1970].

Benvenuti, A., et al. to Wilson, R. R., 14 March 1972 [at least 9 copies were made for authors].

Rubbia, C. and Sulak, L. "Analysis of Neutrino Events with no Associated Muon" [NAL TM, XT, 18 August 1973].

Levedahl, K., Piccioni, R., Rubbia, C., Summers, S., Sulak, L., and Wagner, R. "Hadronic Shower Characteristics and Chamber Efficiencies for Neutrino Events With and Without Muons" [NAL TM, XT, 20 August 1973].

Sulak, L. R. "Evaluation of Muonless Events" [NAL TM, XT, 3 September 1973].

Cline, D. "E1A Detector Modifications for September-December Neutrino Runs" [NAL TM, XT, undated].

Rubbia, C. and Sulak, L. "Possible Improvements in E1A Detector to Augment Muon Acceptance in October Run" [NAL TM, XT, 3 September 1973].

Cline, D. "Performance of Revised E1A Detector for μ-less Event Search" [NAL TM, XT, 1 October 1973].

Cline, D. "Detection of μless Events at the Level of 1%" [NAL TM, XT, 10 October 1973].

Cline, D. "Statistical Analysis of the μless Events from the Test Run" [NAL TM, XT, 11 October 1973].

Aubert, B. "Comments on the Test Run" [NAL TM, XT, 15 October 1973].

Cline, D. "Unified Analysis of the μless Events in the November-December Runs" [NAL TM, XT, 16 October 1973].

Aubert, B. to Gargamelle Collaboration, 16 October 1973.

Rubbia, C. "A Few Thoughts on the New Detector" [NAL TM, XT, 8 November 1973].

Cline, D., Mann, A. K., Reeder, D. D., Rubbia, C. to Lagarrigue, A., 13 November 1973. Unsigned version brought by C. Rubbia to A. Lagarrigue. Signed version in private papers of A. K. Mann. Photocopies in files of many Gargamelle participants.

Imlay, R. and Messing, F. "Beam Associated Muons Entering E1A Detector" [NAL TM, XT, 15 November 1973].

Sulak, L. R. "Early Study of $\bar{\nu}$ Run Nov. '73" [NAL TM, XT, 17 November 1973].

Ford, W. T. and Mann, A. K. "Method to Find R_{corr} (N/C) with Minimum Reliance on Monte Carlo Calculations" [NAL TM, XT, 28 November 1973].

Imlay, R. "Punchthrough" [NAL TM, XT, 29 November 1973].

Imlay, R. "Calculation of Punch Through for the Horn Data" [NAL TM, XT, 29 November 1973].

Cline, D. "Are We Seeing a μless Signal at the Level of 10%?" [NAL TM, XT, 13 December 1973].

Cline, D. "Data Reported at NAL Talk, December 6, 1973." Includes photocopies of transparencies. [NAL TM, XT of cover memorandum with xeroxed manuscript, 13 December 1973].

Aubert, B., Ling, F., and Imlay, R. [Punch Through Calculations] [NAL TM, XT, 18 December 1973].

Mann, A. K. "Summary of Rescan of 300 GeV $\bar{\nu}$ Run for N/C Experiment" [NAL TM, XT and xeroxed spark chamber photographs, 26 January 1974].

Imlay, R. "Calculation of R = $(\nu_\mu \rightarrow \text{no } \mu)/(\nu_\mu \rightarrow \mu)$" [NAL TM, XT, 2 (February?) 1974].

Mann, A. K. [Which Paper to Publish?] [NAL TM, XT, 11 February 1974].

Cline, D. "Does R Grow Rapidly with Energy?" [NAL TM, XT, 12 February 1974].

Cline, D. "Reanalysis of the Test Run" [NAL TM, XT and figure, 22 February 1974].

Sulak, L. R. (?) "Consistency Check of NAL and Harvard Monte Carlo Programs for Experiment 1A" [NAL TM, XT and computer output, 5 March 1974].

Cline, D. "An Elegant, Obvious Solution for R^ν and $R^{\bar{\nu}}$" [NAL TM, XT, 14 March 1974].

Cline, D. "A Simple (and Simple-Minded) Explanation for the High-y Events" [NAL TM, XT and figures, 14 March 1974].

Sulak, L. R. and Rubbia, C. "Probability for a Hadronic Shower to Produce a Muon in the Experiment 1A Magnet" [NAL TM, XT, 18 April 1974].

HARVARD UNIVERSITY

THE GRADUATE SCHOOL OF ARTS AND SCIENCES

VE RI TAS

THESIS ACCEPTANCE CERTIFICATE

(To be placed in Original Copy)

The undersigned, appointed by the

Division

Department

ad hoc Committee for the Ph.D. degree in Physics and History of Science

have examined a thesis entitled: Volume II: "Large Weak Isospin and the W Mass"

presented by Mr. Peter Galison

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature Howard M. Georgi III

Typed name H.M. Georgi, III (Chairman)

Signature J. P. Preskill

Typed name J.P. Preskill

Signature S.L. Glashow

Typed name S.L. Glashow

Date April 13, 1983

LARGE WEAK ISOSPIN AND THE W MASS

A thesis presented

by

Peter Louis Galison

to

The Ad Hoc Committee on Physics and History
of Science in partial fulfillment of
the requirements for the degree
of Doctor of Philosophy in
Physics and History of Science

Part II. Physics

Harvard University
Cambridge, Massachusetts
April, 1983

ABSTRACT

A variation of the usual $SU(2)_L \times U(1)_Y$ weak interaction theory is proposed in which strongly coupled scalars ϕ transform as N N's of $SU(2)_L$ and as an $(N, \bar{N})$ of a global $SU(N)_L \times SU(N)_R$. The gauge group is embedded in the global group so that when $SU(N)_L \times SU(N)_R$ breaks down to $U(N)_V$ it breaks $SU(2)_L \times U(1)_Y$ to $U(1)_{EM}$. This scheme preserves the tree level relation $M_Z/M_W = 1/\cos\theta_W$. Radiative corrections (ΔM_W) to M_W are then discussed and it is found that a) the screening theorem holds, i.e. $\Delta M_W \sim \alpha M_W^2 \ell n\, \xi$ not $\alpha M_H^2 \ell n\, \xi$ where $\xi \equiv \frac{M_H^2}{M_W^2}$ and M_H = scalar mass $\sim$ 1 TeV. A simple symmetry argument accounts for this. b) Radiative corrections increase with the size N of the weak multiplet as N^4. If we demand $M_W^{theor} < (M_W^{CERN} + 3$ standard deviations) then $N < 5$ (M_W^{CERN} is as reported from the UA1 $p\bar{p}$ experiment and we take $\xi \sim \alpha^{-1}$). The model may be interpreted as an effective theory for an underlying renormalizabile technicolor theory.

TABLE OF CONTENTS

PAGE

I. INTRODUCTION 2

A) Motivation 2

B) The Model 3

II. REVIEW OF THE CHIRAL REPRESENTATION OF A SINGLE DOUBLET 9

A) Linear Representation 9

B) Non Linear Representation 12

III. RESULTS FROM TREE LEVEL: MASS SPECTRUM AND SYMMETRIES 17

A) Higgs Potential 17

B) Mass Spectrum 21

IV. RADIATIVE CORRECTIONS TO W-MASSES OF ORDER $M_H^2 \ln \xi$ 25

V. THE GENERALIZED SCREENING THEOREM 34

A) Why There are no $\mathcal{O}(\alpha M_H^2 \ln \xi)$ Corrections to M_W^2 34

B) Illustration of the "Screen Test" . . . 36

VI. RADIATIVE CORRECTIONS TO M_W, M_Z OF ORDER $\alpha M_W^2 \ln \xi$ 38

A) Momentum Dependent Part of Graphs with Two Internal Scalars 39

B) Martian Graphs 41

C) Loops Involving Unphysical Higgs . . . 44

D) Electric Charge Renormalization 48

E) No Ghosts, No Infrared Catastrophe. . . 49

VII. OTHER TERMS IN THE CHIRAL LAGRANGIAN 50

VIII. CONCLUSION 58

FIGURES AND FIGURE CAPTIONS 64

TABLE OF CONTENTS
(continued)

	PAGE
APPENDIX: PGB MASSES	72
REFERENCES	78
ACKNOWLEDGEMENTS	79

INTRODUCTION

I. INTRODUCTION

A. Motivation

The appealing idea has often been advanced that the Higgs scalars, which mediate the standard weak interaction symmetry breaking, are constructed of fundamental fermions. Just as the high energy, renormalizable theory of QCD gives rise to pions so, it is conjectured, might a high energy ($\sim$1 TeV) renormalizable technicolor theory produce (as technipions) the low energy scalar sector of the SU(2) x U(1) electroweak theory. Thus in principle if we knew the high energy theory we could derive the low energy theory by dynamical calculations.

In practice calculation of dynamical symmetry breaking and bound state spectra from first principles in QCD have proven intractable. For many purposes this difficulty has been productively sidestepped by exploiting the symmetries of an effective chiral Lagrangian. The basic idea is this: although we do not know the detailed dynamics of quark binding into bound states, we do know that the QCD Lagrangian has an $SU(3)_L \times SU(3)_R$ global symmetry. The low energy meson spectrum reflects the presence of this symmetry. Our ignorance of the detailed dynamical behavior of the theory is lumped into phenomenological parameters in the effective Lagrangian.

The global symmetry allows many interactions, some renormalizable and others nonrenormalizable. One supposes in the chiral Lagrangian philosophy that all such terms are present, though some will be suppressed by factors of Λ_{CSB} (the SU(3) x SU(3) chiral symmetry breaking scale). Many successful predictions have been made on this basis.

The problem with QTD (quantum technicolor dynamics) is worse than that of QCD. Not only do we lack the calculational tools with which to derive low energy Higgs physics, we don't even have a plausible QTD Lagrangian. However, common to many QTD schemes is an $SU(N)_L \times SU(N)_R$ global symmetry. These arise as well in the composite/fundamental Higgs model of Georgi and Glashow [1] and Georgi and McArthur [2].

In what follows we shall explore some consequences of such an $SU(N)_L \times SU(N)_R$ symmetry for weak interaction phenomenology. In particular, the symmetry will be used here to construct an effective chiral $SU(N)_L \times SU(N)_R$ Lagrangian from which the Z^0 and $W^{\pm}$ masses will be calculated.

B. The Model

Recently Longhitano [3,4] constructed an effective chiral Lagrangian of the $SU(2)_L \times U(1)$ theory. He exploited the fact that the single Higgs doublet of the Weinberg-Salam $SU(2)_L \times U(1)_Y$ model (henceforth the "standard model") could be put in an $SU(2)_L \times SU(2)_R$ chiral representation. In this representation it was straightforward to investigate the consequences of a very massive fundamental Higgs scalar for low energy weak interaction phenomenology. Inter alia Longhitano calculated radiative corrections to the $W^{\pm}$ and Z^{o}.

If one takes seriously the view that Longhitano's Higgs are composite, however, they are not the only important large logarithmic corrections to M_W. In particular, the doublet that gets a vev might be only one of several doublets involving massless scalars. Because the $SU(2)_L \times U(1)_Y$ gauge interactions weakly break the global symmetry, some of the Goldstone bosons (GB's) acquire a mass, and are referred to as Pseudo Goldstone Bosons (PGB's). If the masses of these PGB's are sufficiently large they can induce large logarithmic radiative corrections to M_W through scalar loops.

Peskin and Renken [5] have explored this phenomenon for a variety of technifermion symmetries consistent with the single heavy doublet calculation of Longhitano. (Peskin and Renken allow the global symmetry to be broken by weak and color interactions.) Taken together the Longhitano and Peskin-Renken calculations comprise a self-consistent effective chiral theory calculation.

The present model is the higher weak isospin generalization of the Longhitano-Peskin-Renken work. But instead of supposing the Higgs to occur in a single doublet we take the Higgs to transform as N N-tuplets of $SU(2)_L$. Φ, the Higgs matrix composed of these N N-tuplets, is supposed to transform as an $(N, \bar{N})$ under $SU(N)_L \times SU(N)_R$. The Higgs potential is then constructed of all terms with the $SU(N)_L \times SU(N)_R$ chiral symmetry.

The vacuum expectation value of the Higgs field breaks the $SU(N)_L \times SU(N)_R$ down to $SU(N)_V$. By judiciously embedding the gauged $SU(2)_L \times U(1)_Y$ in the global chiral group, the spontaneous breakdown of the chiral symmetry will break the gauge group down to the $U(1)_{EM}$ of electromagnetism. In particular the breaking preserves the phenomenologically successful tree level relation

$$\frac{M_Z}{M_W} = \frac{1}{\cos\theta_W} \ . \tag{1.1}$$

We enumerate the basic assumptions of the effective $SU(N)_L \times SU(N)_R$ theory to be considered here as follows:

(1) The Higgs come in N fundamental representations of SU(2) each of dimension N. Charge descends from positive on top of each multiplet to negative at the bottom.

These N multiplets are arranged to form an NxN matrix, ϕ, with the neutral fields along the diagonal. We divide ϕ into hermitian and antihermitian parts:

$$\phi = \frac{1}{\sqrt{2}} [\omega X^o + P^\alpha X^\alpha + i(\eta X^o + \pi^\beta X^\beta)] \qquad (1.2)$$

where η, ω, P and π are hermitian fields, X^α are the generators of SU(N) with $\alpha = (1, \ldots, N^2-1)$. $X^o_{ij} = (1/\sqrt{N})\ \delta_{ij}$; $\mathrm{tr}\,(X^A X^B) = \delta_{AB}$ $(A, B = 0, \ldots, N^2-1)$.

(2) Because of their weak-isospin representation content, the scalars do not couple to quarks and leptons. They serve only to give mass to the gauge bosons. For simplicity of exposition we assume they are the only source of gauge boson masses. With no difficulty it could be assumed that they exist in addition to the standard model Higgs doublet, which gives mass through Yukawa couplings to the quarks and leptons.

(3) ϕ transforms as N N's of $SU(2)_L$, whose generators are normalized according to $\mathrm{Tr}(T^a T^b) = \frac{N(N^2-1)}{12}\,\delta_{ab}$. (1.3)

(4) The Lagrangian is characterized by a $SU(N)_L \times SU(N)_R$ global symmetry under which ϕ transforms as an $(N, \bar{N})$. This symmetry is imposed on $V(\phi)$ by fiat. This is unlike

the $SU(2)_L \times SU(2)_R$ symmetry in the standard model which, follows "accidentally" in the chiral representation of the $SU(2)_L \times U(1)_Y$ gauge symmetry, with just one $SU(2)_L$ doublet. By adding a determinent term to the scalar potential we restrict the symmetry to $SU(N)_L \times SU(N)_R$. The Lagrangian $\mathscr{L}$ is given by

$$\mathscr{L} = \mathrm{tr}\left[\left(\partial\phi - ig(T^1W^1\phi + T^2W^2\phi + T^3W^3\phi) + ig'B\phi T^3\right)(\mathrm{herm.conj.})\right] + \mathscr{L}_G + V(\phi) + (\text{other chiral terms}). \tag{1.4}$$

in which $e = gs = g'c$ with s and c sine and cosine of the weak mixing angle θ_W. $\mathscr{L}_G$ includes the gauge fixing term and the ghost term both of which will be discussed in Section VI. Finally $V(\phi)$ is the effective potential which includes all possible $SU(N)_L \times SU(N)_R$ invariant terms that are consistent with the $SU(2)_L \times U(1)_Y$ gauge symmetry.

In this paper we explore the radiative corrections to M_Z and M_W using this model. Sections II and III are devoted to the symmetries and mass spectra of the tree level theory. In particular it is shown that the tree relation $M_Z/M_W = 1/\cos\theta_W$ holds for any N, with N the dimension of the scalar multiplet [6]. Next, in Section IV it is shown that the model obeys the "screening theorem", i.e., first-order radiative corrections to the vector boson masses are proportional to αM_W^2 rather than to αM_H^2 (Higgs mass squared). This has been observed in the standard model by Veltman [7], Sirlin [8], Longhitano [3,4], and others. In Section V we offer a simple way of understanding the screening theorem and provide an illustrative toy model which shows how, in another context, the screening theorem might not apply. Then in Section VI it is demonstrated by explicit calculation that the radiative corrections to M_Z^2 and $M_{W^\pm}^2$ are of order $\alpha N^4 M_W^2 \ln \xi$ where N is the dimension of the Higgs scalar multiplets and $\xi = M_H^2/M_Z^2$. (We will neglect the difference between $\log M_H^2/M_W^2$ and $\log M_H^2/M_Z^2$.)

The second part of our calculation (Section VII) involves the examination of other chiral symmetry breaking terms that appear in the Lagrangian. These contributions are most easily treated in the non-linear chiral representation of the scalars in which the heavy particles do not appear. Using this formalism we find that the heaviest of the PGB's get a mass of order αM_W. These are too small by a factor of α to contribute significantly to our calculation. Finally it is shown that the symmetry breaking terms involving more than four covariant derivatives are also suppressed relative to those calculated in Sections V and VI by powers of α. For completeness the connection between the linear and non-linear representations is discussed in some detail.

The last section, VIII, brings together all of the radiative corrections to M_W as a function of N, the dimension of the scalar multiplet. An experimental determination of the vector boson masses would then specify the maximum size of N.

II. REVIEW OF THE CHIRAL REPRESENTATION OF A SINGLE SCALAR DOUBLET

A. Linear Representation

Several features of the $SU(N)_L \times SU(N)_R$ theory are usefully highlighted by comparing them with their analogues in the standard model with just one Higgs doublet. In particular we contrast the linear and non linear representations. If ϕ is the SU(2) doublet $\begin{pmatrix}\phi_o\\ \phi_-\end{pmatrix}$ then $\delta\phi = i\varepsilon^a T^a \phi$, (a = 1 - 3). Since the doublet representation of $SU(2)_L$ is pseudoreal, there exists a real antisymmetric matrix C with $C^2 = 1$ such that $CT + T^*C = 0$. By convention T^1 and T^3 are real and T^2 is imaginary. Therefore $[C, T^2] = \{C, T^{1(3)}\} = 0$ and so $C\phi^* = \tilde{\phi} = iT^2\phi$ transforms as $\delta\phi = i\varepsilon^a T^a \tilde{\phi}$. Therefore $M_{ij} \equiv \phi_i\, \delta_{j1} + \tilde{\phi}_i \delta_{j2}$ transforms as does ϕ under $SU(2)_L \times U(1)_Y$. It is useful to write

$$M = \frac{1}{\sqrt{2}}\,(\sigma + iT^a\pi^a). \qquad (2.1)$$

Then the Higgs potential

$$V(M) = \mu^2\, tr(MM^+) + \lambda[tr(MM^+)]^2 \qquad (2.2)$$

can be rewritten as

$$V(M) = \lambda\ (tr[MM^+ - v^2])^2 \qquad (2.3)$$

with $v^2 = -\mu^2/2$. Therefore $\langle M \rangle = v\delta_{ij}$.

We see that the gauged $SU(2)_L \times U(1)_Y$ automatically implies that the global $SU(2)_L \times SU(2)_R$ be a symmetry of the scalar sector. If only $SU(2)_L$ were gauged then the kinetic term

$$tr\ \{\mathscr{D}M(\mathscr{D}M)^+\}, \qquad (2.4)$$

where $\mathscr{D}_\mu M \equiv \partial_\mu M - igT^a W_\mu^a M$ would also be invariant under $SU(2)_L \times SU(2)_R$. This is because when $M \rightarrow LMR^+$, $\mathscr{D}M \rightarrow L(\mathscr{D}M)R^+$. When the $U(1)_Y$ is also gauged, however, then

$$\mathscr{D}_\mu M \equiv \partial_\mu M - igT^a W_\mu^a M + ig'B_\mu MT^3,\ (a=1-3) \qquad (2.5)$$

and the kinetic term does not have the chiral $SU(2)_L \times SU(2)_R$ symmetry.

To calculate perturbatively we expand the field M around the vacuum, by defining $M = M' + \langle M \rangle$, where $\langle M \rangle$ is the vacuum expectation value. This breaks the symmetry. Dropping the primes

$$V(M) = \lambda[tr(MM^+ + v(M + M^+)]^2. \qquad (2.6)$$

Thus only the $SU(2)_V$ symmetry, i.e. $M \to e^{i\vec{\varepsilon}\vec{T}} M e^{-i\vec{\varepsilon}\vec{T}}$ survives. If we think of the original generators of the chiral symmetry as the $SU(2)_L$ and the $SU(2)_V$ generators then we see that the $SU(2)_L$ generators are broken. Since these are gauged symmetries, the corresponding GB's are eaten to give massive gauge W's. Eq. (2.6) makes it clear that the GB's come from the antihermitian piece of M and that the hermitian piece gets a mass $(2\lambda f^2)^{1/2}$.

Expanding the kinetic term $\mathrm{tr}\{\mathcal{D}M(\mathcal{D}M)^+\}$ leads straightforwardly to the tree level gauge mass matrix for the gauge bosons:

$$M^2_{W_i W_j} = 2g^2 v^2 \delta_{ij} \text{ and}$$

$$M^2_B = 2g'^2 v^2. \tag{2.7}$$

The W^3 - B (neutral) mass submatrix is thus of the form,

$$\chi^2_{neutral} = 2v^2 \begin{array}{c} \begin{matrix} W^3 & \quad B \end{matrix} \\ \begin{pmatrix} g^2 & a \\ a & g'^2 \end{pmatrix} \end{array} \tag{2.8}$$

for some a which we now determine. Electric charge conservation implies that there is an unbroken electromagnetic U(1) symmetry which prohibits a photon mass term. The masslessness of the photon implies that det $\chi^2 = 0$ thus $a = \pm gg'$ which gives us the non-zero mass eigenvalue,

$$M_Z^2 = gg'/(g^2 + g'^2). \tag{2.9}$$

Consequently,

$$M_Z/M_W = 1/\cos\theta_W$$

since conventionally

$$e = g'c = gs$$

with

$$c \equiv \cos\theta_W \text{ and } s \equiv \sin\theta_W. \tag{2.10}$$

In short the surviving $SU(2)_V$ symmetry together with electric charge conservation yield the natural relation $M_Z/M_W = 1/\cos\theta_W$.

B. Non-Linear Representation

There is a non linear representation of the scalars which is frequently more convenient than the linear representation described thus far. Supposing the σ field to be heavy (of order 1 TeV) one can simply calculate with σ present, throwing away contributions to the W mass either suppressed by additional factors of α or lacking the $\log\,(M_H^2/M_W^2)$ enhancement. It is also possible to dispose of the heavy field from the outset. In eq. (2.2) if we let $\mu^2 \to \infty$ the potential becomes infinitely steep at the minimum, cutting off quantum fluctuations in certain directions of M. From eq. (2.3) the minimum of V(M) occurs when

$$MM^+ = v^2. \qquad (2.11)$$

Since $MM^+ = \sigma^2 + \vec{\pi}^2$ this implies that

$$MM^+ = \vec{\pi}^2 + \sigma^2 = v^2 \qquad (2.12)$$

If we now define $\Sigma(x) = M(x)/v$ then

$$\Sigma\Sigma^+ = 1 \qquad (2.13)$$

at the minimum. Effectively, the σ field has been removed from the theory since it can be reparameterized as $\sqrt{1 - \pi(x)^2}$. Alternatively we simply write

$$\Sigma = \exp i\underset{\sim}{\pi}/f \quad \text{with } \Sigma \to U_L \Sigma U_R^+, \qquad (2.14)$$

$\underset{\sim}{\pi} = \pi^\alpha T^\alpha$, $T^\alpha = \tau^\alpha/2$, U_L and U_R^+ are independent left and right SU(2) generators, and f is of the order of 1 TeV. This is a convenient notation which will prove useful in the SU(N) case.

We can therefore write an effective non linear Lagrangian, $\mathcal{L}^{NL}_{eff}$ as

$$\mathcal{L}^{NL}_{eff} = f^2 \, \mathrm{tr}[(\mathcal{D}_\mu \Sigma)(\mathcal{D}^\mu \Sigma)^+] + \mathcal{L}_G$$

$$+ \text{(other chiral terms)} \, . \qquad (2.15)$$

$\mathcal{L}_G$ includes the gauge fixing term, the Fadeev-Popov ghost term, and the kinetic terms for the gauge fields.

Obviously when we actually calculate M_W, M_{Z^o} to $\mathcal{O}(\alpha)$, the masses must be the same for both linear and non-linear representations. The large logarithmic enhancement $\ln \xi \equiv \ln(M_H^2/M_W^2)$ enters, however, in different ways in the two cases. In the linear case the Higgs mass enters through massive scalar propagators into the vacuum polarization diagrams. When these are renormalized one gets a correction to the mass of

$$\Delta M_W^2 \sim \alpha \, p^2 \, \ln[(p^2 + M_H^2)/\mu^2] \qquad (2.16)$$

which for $p^2 = \mu^2 \doteq M_W^2$ gives the $\ln \xi$ enhancement. In the non linear representation, by contrast, no heavy scalars are present. We may nonetheless divide the chiral perturbation 1-loop integrals in two parts. Schematically

$$\Delta M_W^2 \sim \alpha p^2 \int_o^\infty \frac{d^4k}{k^2(k+p)^2} = \alpha p^2 \int_o^{\Lambda_{CSB}} \frac{d^4k}{k^2(k+p)^2}$$

$$+ \alpha p^2 \int_{\Lambda_{CSB}}^\infty \frac{d^4k}{k^2(k+p)^2} \qquad (2.17)$$

where Λ_{CSB} is the 1-TeV chiral symmetry breaking scale. In the second term of the r.h.s. the underlying technifermions can no longer be considered as condensed into chiral symmetry breaking technipions. Since chirally symmetric contributions to the vacuum polarization do not break $SU(2)_L \times U(1)_Y$, the second integral does not renormalize the mass of the vector bosons. Thus in the chiral theory the mass correction is

$$\Delta M_W^2 \sim \alpha\ p^2\ \ell n(\Lambda_{CSB}^2/p^2) \sim \alpha\ M_W^2\ \ell n\ (\Lambda_{CSB}^2/M_W^2) \qquad (2.18)$$

and this is also of the form $\alpha\ M_W^2\ \ell n\ \xi$. The chiral symmetry breaking scale Λ_{CSB} thus serves as the regulator of the theory.

In this paper we first exploit the linear representation of the chiral theory (massive fields present) to calculate radiative corrections to M_W, M_{Z^o} from vertices in $tr[\mathscr{D}\phi(\mathscr{D}\phi)^+]$. This has the advantage of exhibiting clearly the relation of terms in the effective Lagrangian to Feynman diagrams. When we turn to a discussion of other terms in $\mathscr{L}_{eff}$ it is easier to argue directly from the non-linear chiral representation (massive fields absent.) These contributions will be shown to be small compared to those from vertices in $\mathrm{tr}\{\mathscr{D}\phi(\mathscr{D}\phi)^{+}\}$.

III. RESULTS FROM TREE LEVEL: MASS SPECTRUM AND SYMMETRIES

A. Higgs Potential

Proceeding by analogy with the SU(2) case we define

$$\phi = \frac{1}{\sqrt{2}} \left(P^\alpha X^\alpha + \omega X^o + i \left[\pi^\beta X^\beta + \eta X^o\right]\right) \tag{1.2}$$

with X^α the generators of SU(N) normalized as in (1.1) such that $tr(X^\alpha X^\beta) = \delta_{\alpha\beta}$. The linear representation of the scalar fields allows easy calculation with the dimension four term, $tr[\mathcal{D}\phi(\mathcal{D}\phi)^+]$, and the dimension four terms in the potential. In addition, initially including the massive scalars will clarify the connection between the linear representation (which has both light and heavy scalars) and the nonlinear representations (which involves only light scalars).

What renormalizable $SU(N)_L \times SU(N)_R$ symmetric and $\phi \to -\phi$ invariant potential terms exist? Since ϕ is an $(N, \bar{N})$, it transforms linearly, as

$$\phi \to e^{i\lambda_L^a X^a} \phi\, e^{-i\lambda_R^b X^b},$$

where

$$a \text{ and } b = 0, \ldots, N^2-1. \tag{3.1}$$

This implies that up to a constant

$$V(\phi) = \mu^2 \, tr(\phi\phi^+) + A \, tr(\phi\phi^+)^2 + B[tr(\phi^+\phi)]^2 \qquad (3.2)$$

is the most general $V(\phi)$ satisfying our criteria. Positivity requires $A + B > 0$ and $A + BN > 0$.

If $B < 0$ it is convenient to rewrite Eq. (3.2) as,

$$V(\phi) = - BN \, tr\{(\phi\phi^+) - \frac{1}{N} [tr(\phi\phi^+)]\}^2$$

$$+ (A + BN) \, tr(\phi\phi^+ - v^2)^2 \qquad (3.3)$$

with

$$v^2 \equiv - \mu^2/2(A + BN) \qquad (3.4)$$

The term in curly brackets is manifestly positive. Since both terms are positive definite, if we find a $<\phi>$ such that $V(<\phi>) = 0$ then $<\phi>$ is a minimum. Such a $<\phi>$ is:

$$<\phi> = v \, \delta_{ij} \; (i,j = 1\text{-}N). \qquad (3.5)$$

For $A > 0$, $B > 0$ eq. (3.2) can be written as the sum of two positive definite terms,

$$V(\phi) = A \, tr(\phi\phi^+ - v^2)^2 + B[tr(\phi\phi^+ - v^2)]^2, \qquad (3.6)$$

with the same v as above and $\langle\phi\rangle = v\delta_{ij}$ is again a minimum. Notice that $U(N)_V$ (and in particular the $SU(2)_V$ embedded in the $U(N)_V$) is left unbroken when takes its vacuum expectation value. As we will see it is this remaining symmetry that is responsible for preserving the relation

$$\frac{M_Z}{M_W} = \frac{1}{c}$$

at tree level.

In the unbroken theory we assume that the gauged $SU(2)_L$ is contained in the the $SU(N)_L$ and that the $U(1)_Y$ of hypercharge is contained in the $SU(N)_R$. When ϕ takes its vev, $\langle\phi\rangle$ we express the $V(\phi)$ of eq. (3.6) in terms of the shifted field:

$$V_{broken}(\phi) = A\ tr[\phi\phi^+ + v(\phi + \phi^+)]^2$$

$$+ B[tr(\phi\phi^+ + v(\phi + \phi^+))]^2 \qquad (3.7)$$

The $SU(N)_L \times SU(N)_R$ symmetry has broken to $U(N)_V$ and in doing so has broken $SU(2)_L \times U(1)_Y$. One combination of the $SU(2)_L$ and $U(1)_Y$ generators survives in the unbroken $U(N)_V$,

$$\phi \to e^{i\lambda T^3}\ \phi\ e^{-i\lambda T^3}. \qquad (3.8)$$

We identify this U(1) with electromagnetism.

For completeness we finish our discussion of Higgs minima by showing that the choice of $A < 0$, $B > 0$ leads to a vev that violates the custodial $SU(2)_V$. Under this choice of parameters $V(\phi)$ (Eq. 3.2) can be written,

$$V(\phi) = - A\{[\mathrm{tr}(\phi\phi^+)]^2 - \mathrm{tr}(\phi\phi^+)^2\}$$

$$+ (A + B)\ [\mathrm{tr}(\phi\phi^+ - v'^2)]^2. \qquad (3.9)$$

The term in curly brackets is positive definite by the Schwartz inequality. Therefore again $V(\phi)$ has a minimum at $<\phi>$ if $V(<\phi>) = 0$. Such a $<\phi>$ is:

$$<\phi> = \sqrt{N}\ v'\ \delta_{1j}\ \delta_{1i} \qquad i,j = 1, \ldots, N \qquad (3.10)$$

where

$$v' \equiv \mu^2/2N(A + B). \qquad (3.11)$$

Clearly the custodial $SU(2)_V$ symmetry is violated in the broken tree level Lagrangian for this choice of $<\phi>$. Consequently the natural relation $M_Z/M_W = 1/c$ does not hold. We therefore exclude $A < 0$ $B > 0$ on the grounds of incompatibility with neutral current experiments. All other choices of A,B preserve the natural relation.

B. Mass Spectrum

To find the mass spectrum of the theory we identify the coefficients of $\phi\phi^+$ in eq. (3.7). These are included in

$$V(\phi)_{quad} = - Av^2 \, tr(\phi^+ + \phi)^2 - Bv^2[tr(\phi^+ + \phi)]^2. \quad (3.12)$$

From the definition of ϕ in eq. (3.2), we see that

$$(\phi^+ + \phi) = \sqrt{2}\, (P^\alpha X^\alpha + \omega X^o) \quad (3.13)$$

so only the hermitian fields P^α and ω get a mass. The antihermitian part of ϕ, $(\phi - \phi^+) = \sqrt{2}\, i(\pi^\alpha X^\alpha + \eta X^o)$ represents the N^2 Goldstone bosons. They will receive detailed consideration in sections VI and VII.

After symmetry breakdown $\mathcal{L}^L_{eff}$ retains a $U(N)_V$ symmetry. Under this symmetry ϕ transforms as an N^2-1 and as a singlet. Thus we expect N^2-1 particles of mass M_P, 1 particle of mass M_ω, and N^2 massless particles from the N^2-1 π's and the single η.

Explicitly, the quadratic piece of eq. (3.12) is:

$$V(\phi)_{quad} = - 2Av^2 \, tr\{(P^\alpha X^\alpha + \omega X^o)\, (P^\beta X^\beta + \omega X^o)\}$$

$$- 2Bv^2 \, [tr(P^\alpha X^\alpha + \omega X^o)]^2. \quad (3.14)$$

Expanding $V(\phi)_{quad}$ and using the normalization $tr(X^a X^b) = \delta_{ab}$ we can pick out the coefficients of terms quadratic in fields. One finds that,

$$M_P^2 = 4Av^2$$

$$M_\omega^2 = 4v^2 \ (A + BN) \text{ and}$$

$$M_\pi^2 = M_\eta^2 = 0. \tag{3.15}$$

In the absence of an A term in the potential (eq.3.6) the ω "trace particle" is the only massive particle and there are $2N^2$-1 PGB's. This follows from the B term's $0(2N^2)$ symmetry which spontaneously breaks down to $0(2N^2-1)$. (The symmetry of the A term is $U(N)_L \times U(N)_R$ which breaks down to $U(N)_V$.)

Finally we may wish to leave only $SU(N)_V$ (and not $U(N)_V$.) as the surviving symmetry of the broken Lagrangian. This can be done by adding a term

$$C \det(\phi^+ - \phi)^2 \tag{3.16}$$

to the effective theory. Such a term gives mass to the η particle only. C by assumption sets the mass of η at $\sim$1 TeV.

Our Lagrangian eq. (1.4) also includes a kinetic term $tr(\mathcal{D}\phi(\mathcal{D}\phi)^+)$:

$$tr\{[\partial_\mu\phi - igT^a W^a_\mu\phi + ig'B_\mu\ \phi T^3][\partial^\mu\phi + ig\phi^+ T^b W^\mu$$

$$- ig'B^\mu T^3\ \phi^+]\}. \tag{3.17}$$

By convention we take $T^a \equiv$ generators of SU(2) in the spin-p representation (N = 2p+1); $T^\pm = T^1 \pm iT^2$; $tr(T^+T^-)/2 = tr(T^3T^3) = N(N^2-1)/12$; $W^\pm = (W^1 \pm iW^2/\sqrt{2}$. Using our definition (eq. 1.2)

$$\phi \equiv \frac{1}{\sqrt{2}}\,[\omega X^o + p^\alpha X^\alpha + i\ (\pi^\beta X^\beta + \eta X^o)]$$

we shift the ϕ field in eq. (3.17) to deduce the W gauge boson mass matrix:

$$M^2_{W_a W_b} = 2g^2 v^2\ tr(T^a T^b) = 2g^2 v^2\ tr(T^3T^3)\delta_{ab}$$

$$M^2_{BB} = 2g'^2 v^2\ tr(T^3T^3). \tag{3.18}$$

Exactly as in the standard Weinberg-Salam case the surviving $SU(2)_V$ of the W part of the mass matrix and charge conservation implies the natural relation $M_Z/M_W = 1/\cos\theta_W$. Separately the tree level masses of the Z^o and W are the same as the Weinberg-Salam case except they are multiplied by the normalization factor $tr(T^3T^3) \equiv \Delta$:

$$M_Z^2 = \frac{2g^2v^2\Delta}{c^2}$$

$$M_W^2 = 2g^2v^2\Delta \ . \qquad (3.19)$$

IV. RADIATIVE CORRECTIONS TO W-MASSES OF ORDER $\alpha M_H^2 \ln \xi$

Radiative corrections to the $SU(2)_L \times U(1)_Y$ theory have been studied in a simple renormalization framework by Sirlin [8]. We will follow his scheme in which for $c \equiv \cos\theta_W$, $s \equiv \sin\theta_W$:

$$\frac{M_Z(\text{physical})}{M_W(\text{physical})} = \frac{1}{c} \tag{4.1a}$$

$$M_W(\text{tree}) = \left(\frac{\pi\ \alpha}{\sqrt{2}\ G_\mu s^2}\right)^{1/2} = 77.9\ \text{GeV} \tag{4.1b}$$

$\alpha = 137^{-1}$ is defined through the emission of a zero momentum photon from a charged particle. (4.1c)

$G_\mu = 1.17 \times 10^{-5}\ \text{GeV}^{-2}$ is the quantity that appears in the effective local V-A Lagrangian for $\mu \to \nu_\mu\ e\ \bar{\nu}_e$. In lowest order G_μ fixes the vev, $v \equiv \langle\phi\rangle$ by

$$G_\mu = \frac{\sqrt{2}\ g^2}{8\ M_W^2} = \frac{1}{8\sqrt{2}\ v^2\Delta} \tag{4.1d}$$

where eq. (3.19) gives us $M_W^2 = 2v^2g^2\Delta$.

$$s^2 = 0.229 \tag{4.1e}$$

is defined through low energy elastic neutrino scattering processes involving ratios of neutral to charged currents. In this definition of s^2, the weak mixing angle is not changed by radiative corrections.

In this scheme (though not in others) radiative corrections determine the relation between M_W(tree) = 77.9 GeV and M_W(physical). To find these corrections it is useful to define the regularized but unrenormalized self energies by

$$\pi^{\mu\nu}_{WW}(p^2) = A_{WW}(p^2)g^{\mu\nu} + B_{WW}(p^2)p^\mu p^\nu,$$

$$\pi^{\mu\nu}_{ZZ}(p^2) = A_{ZZ}(p^2)g^{\mu\nu} + B_{ZZ}(p^2)p^\mu p^\nu \tag{4.2}$$

in which $\pi^{\mu\nu}_{WW}$ is defined as the amputated self-energy Feynman diagram shown in Fig. 4.1. $\pi^{\mu\nu}_{ZZ}$ is defined similarly.

In the analysis one can isolate and retain only pieces enhanced by a large logarithm $\ell n\ \xi$ where $\xi = M_H^2/M_Z^2$. (Typically we will take $\xi \sim \alpha^{-1}$ to characterize the large mass, M_H.) From Sirlin's analysis the radiative corrections with large logs can be expressed in terms of A_{WW} and A_{ZZ} as:

$$M_W(\text{physical}) = (1 + \frac{\Delta r^{\text{other}}}{2} + \Delta r^H/2)\ M_W(\text{tree}) \tag{4.3}$$

where

$$\Delta r^H = - \frac{c^2}{s^2 M_W^2} [c^2 A_{ZZ}(-M_Z^2) - A_{WW}(-M_W^2)]$$

$$- \frac{1}{M_W^2} [A_{WW}(-M_W^2) - A_{WW}(0)] - \pi^H_{\gamma\gamma}(0). \qquad (4.4)$$

$\pi^H_{\gamma\gamma}(0)$ is the coefficient of $(p^2 g_{\mu\nu} - p_\mu p_\nu)$ from scalar corrections to the photon propagator. (The minus signs in front of all the terms come from the rotation into Euclidean space of $g^{\mu\nu}$; $p^2_{\text{euclidean}} = - M^2_{W(Z)}$.)

Note that the second term of Sirlin's formula vanishes for any piece of A_{WW} that is independent of momentum. The bracket in the first term of Sirlin's formula vanishes in the limit where the weak mixing is turned off, i.e. when $c^2 \to 1$.

Δr^{other} includes all non-scalar corrections to the self energy diagrams such as lepton and hadron loops. Δr^{other} also includes a bosonic contribution that is independent of the Higgs mass [8] (though this is only $\Delta r/2 = 0.001$). In all

$$\frac{\Delta r^{\text{other}}}{2} = 0.034. \qquad (4.5)$$

Finally for the standard Weinberg-Salam model with a single heavy Higgs doublet of mass $M_H^2 \sim \alpha^{-1} M_Z^2$, there is a contribution of

$$\frac{\Delta r^H}{2} \text{ (1 doublet)} = \left(\frac{11}{24}\right) \frac{\alpha}{4\pi s^2} \ln \xi = 0.006 \qquad (4.6)$$

with $\xi \sim 137$. Thus according to Sirlin [8]:

$$M_W(\text{physical, 1-doublet}) = (1.040)\ 77.9 = 81.0 \text{ GeV.} \qquad (4.7)$$

More recent work by Marciano and Sirlin [9] suggests that the radiatively corrected M_W in the standard model with one heavy doublet should be:

$$\begin{aligned} M_W(\text{physical, 1-doublet}) &= 82 \pm 2.4 \text{ GeV} \\ &= M_W(\text{tree}) \left(1.006 + \frac{\Delta r^{\text{other}}}{2}\right). \qquad (4.8) \end{aligned}$$

Our task is to calculate Δr^H (and therefore M_W) as a function of N using the scalar sector just described.

What are the largest contributions to Δr^H? At first blush loops involving Higgs scalars might appear to give rise to a leading W-mass correction of the order,

$$\Delta M_W^2 \sim \frac{\alpha\, M_H^2}{(4\pi)^2} \ln \xi. \qquad (4.9)$$

The logarithm and $(4\pi)^{-2}$ come from the loop integral. One would expect one such contribution from each of the N Higgs multiplets. A further factor of Δ ($\Delta \equiv tr(T^3T^3)$) comes from the gauge couplings at the vertices. Thus naively we might expect:

$$M_W^2(\text{physical}) = M_W^2(\text{tree}) + \alpha N^4 \Delta \frac{M_H^2}{(4\pi)^2} \ln \xi. \qquad (4.10)$$

However we will see that the contributions of this order (from diagrams (a) - (c) below) conspire to cancel. In this section we exhibit this cancellation by explicit calculation. In section V this cancellation will emerge as a consequence of a suitably generalized version of the "screening theorem", which assures the screening out of the $\mathcal{O}(\alpha M_H^2 \ln \xi)$ radiative corrections.

We now examine the diagrams in Fig. 4.2, each of which includes terms proportional to αM_H^2. Each diagram involves a trace over several fields; fortunately some group identities make their evaluation simple. With the group factors in hand, the standard momentum integrals of these diagrams can then easily be evaluated by dimensional regularization.

As an example of how the group factors are calculated, consider the vertex $W^1 P^\alpha P^\beta$, which is extracted from one of the terms in $\mathrm{tr}[\mathcal{D}\phi(\mathcal{D}\phi)^+]$, using eq. (1.4):

$$\frac{1}{2}\,\mathrm{tr}[(\partial_\mu P^\alpha X^\alpha - i\beta c\; T^1 W^1 P^\alpha X^\alpha)(\partial^\mu(P^\beta X^\beta + i\beta c\; P^\beta X^\beta T^1 W^{1\mu}))], \tag{4.11}$$

where $\beta \equiv e/sc$. It follows that the graph in Fig. 4.2a for two P particles in the loop is equal to:

$$\left(\frac{\beta^2\, c^2\, N\, \Delta}{4}\right) F(M_P M_P) \tag{4.12}$$

where $F(M_P, M_P)$ designates the standard momentum integral of Fig. 4a,b with two P-masses in the scalar propagators.

The product of the two traces is best evaluated using the SU(N) identity:

$$(X^{\alpha})^{a}_{b}(X^{\alpha})^{c}_{d} = \delta^{a}_{d}\delta^{c}_{b} - \frac{1}{N}\,\delta^{a}_{b}\delta^{c}_{d}\ . \qquad (4.13)$$

An almost identical argument leads to an anticommutator in the vertex factor for the $W^1 P^\alpha \pi^\beta$ vertex because one field is hermitian and one antihermitian. Again making use of the group identity (4.13) to find the product of the traces, we find that the graph of type 4.1b with an external W and internal π^β and P^α is:

$$\frac{\beta^2 c^2}{4}\, \mathrm{tr}(\{X^\alpha,X^\beta\}T^1)\,\mathrm{tr}(\{X^\beta,X^\alpha\}T^1)\,F(M_P M_\pi)$$
$$= \frac{\beta^2 c^2}{2}\,(N - 4/N)\,F(M_P,M_\pi)\ . \qquad (4.14)$$

$F(M_1,M_2)$ is defined to be the coefficient of $g_{\mu\nu}$ in the standard 2-scalar loop diagram depicted in Figs. 4.1a, 4.1b. The dimensionally regularized 2-scalar integral has the form:

$$\int [\mu^{2-\omega}]^2 \frac{d^{2\omega}\ell}{(2\pi)^{2\omega}} \left[\frac{(2\ell-p)_\mu (2\ell-p)_\nu}{[(\ell-p)^2 + M_1^2](\ell^2+M_2^2)}\right] \qquad (4.15)$$

in Euclidean space with loop momentum ℓ_μ, external momentum p_μ, and scalar masses M_1 and M_2. After subtraction using the $\overline{\mathrm{MS}}$ renormalization prescription at $p^2 = -M^2_{W(Z)}$ and renormalization point $\mu^2 - M_W^2$, we have for a coefficient of $g_{\mu\nu}$ the following:

$$F(M_1M_2) = 2g_{\mu\nu}\left(\frac{1}{4\pi}\right)^2\left[\int_0^1 dx\left\{(M_2^2 + x(M_1^2-M_2^2) + p^2x(1-x))\right.\right.$$
$$\left.\times\ \ell n\left(\frac{M_2^2 + x(M_1^2-M_2^2) + p^2x(1-x)}{\mu^2}\right)\right\}$$
$$\left. - \int_0^1 (M_2^2 + x(M_1^2-M_2^2) + p^2x(1-x))dx\right] . \qquad (4.16)$$

In all that follows we will only be interested in the portion of F with large logarithms, i.e. logs of $\xi \equiv M_H^2/M_Z^2$. These large logs arise in the linear representation for $M_\eta^2 \sim M_p^2 \sim M_\omega^2 \sim 1$ TeV $\equiv M_H^2$, $\mu^2 = M_W^2$; in the nonlinear representation they come from ratios of $p^2 = -M_W^2$ to $\Lambda_{CSB}^2 \sim 1$ TeV with all scalar masses zero.

It is useful to divide $F(M_1M_2)$ into two pieces:

$$F(M_1M_2) = S(M_1M_2) + S'(M_1M_2) \qquad (4.17)$$

where $S(M_1M_2)$ is the piece proportional to M_H^2 and $S'(M_1M_2)$ is the piece proportional to p^2.

Finally it is useful to have before us the standard integral S(M) for the tadpole diagram in Fig. 4.1c:

$$S(M) = \int \frac{d^{2\omega}\ell}{(2\pi)^2}\,\frac{1}{[(\ell-p)^2 + M^2]} = -\frac{1}{(4\pi)^2}M^2\,\ell n\,\xi \ . \qquad (4.18)$$

To find the contribution of each type of graph in Figs. 4.2a-c we multiply the appropriate standard integral $S(M_1M_2)$ by the group, combinatoric, and coupling constants. Retaining only the contributions to A_{WW} and A_{ZZ} that are proportional to M_H^2 we have the results shown in Table I.

Feynman Diagrams of Type Shown in Fig. 4.2 with External W and the Following Virtual Particles	Contribution to A_{WW} Proportional to M_H^2
$P\pi$	$\frac{\Delta\beta^2 c^2}{2}(N-4/N)[\Sigma]$
PP	$\frac{\Delta\beta^2 c^2 N}{4}[2\Sigma]$
$P\eta$ or $\pi\omega$	$(\Delta\beta^2 c^2/N)[\Sigma]$
$\eta\eta$ or $\omega\omega$ or $P\omega$ or $\pi\eta$ or $\pi\pi$	0
η or ω	$(\Delta\beta^2 c^2/N)[-\Sigma]$
P	$\Delta\beta^2 c^2(N-1/N)[-\Sigma]$

Fig. 4.2 with External Z and the Following Virtual Particles	Contribution to A_{ZZ} Proportional to M_H^2
$P\pi$	$\frac{\Delta\beta^2}{2}(N-4/N)[\Sigma]$
PP	$\frac{\Delta\beta^2 N}{4}(1+4s^4-4s^2)[2\Sigma]$
$P\eta$ or $\pi\omega$	$(\Delta\beta^2/N)[\Sigma]$
$\eta\eta$ or $\omega\omega$ or $P\omega$ or $\pi\eta$ or $\pi\pi$	0
η or ω	$(\Delta\beta^2/N)[-\Sigma]$
P	$\Delta\beta^2\left(N(1-2s^2+2s^4)-1/N\right)[-\Sigma]$

Table I. Contributions to A_{WW} and A_{ZZ} proportional to M_H^2 from Feynman graphs of type shown in Fig. 4.2. $\Delta \equiv tr(T^3T^3)$; $\beta \equiv e/sc$; $\Sigma \equiv (1/4\pi)^2 M_H^2 \ln\xi$.

Since these group, combinatoric and coupling factors will be useful to us later, we have isolated the standard integrals in square brackets. Note that the M_H^2 contributions to A_{WW} and A_{ZZ} involving two internal π's or 4.2c with one π vanish because the mass of the π is zero. By contrast the terms involving internal $\eta\eta$, $\omega\omega$, $P\omega$ or $\pi\eta$ vanish because they involve vanishing commutators from the group factors. When $SU(2)_V$ symmetry is restored by setting $g' = 0$ (i.e. holding e/s constant while $s \to 0$) we see that the corresponding contributions of A_{WW} and A_{ZZ} become equal in Table I as we would expect.

Using the group factors of Table I we can exploit Sirlin's eq. (4.4) to find radiative corrections to M_W and M_Z that are proportional to αM_H^2. Now the relevant contributions to Δr^H come from terms in A_{WW}, A_{ZZ} that are proportional to αM_H^2. These are momentum independent. Thus Sirlin's formula reduces to:

$$\Delta r^H(\propto \alpha M_H^2) = - \frac{c^2}{M_W^2 s^2} [c^2 A_{ZZ} (M_H^2 - A_{WW}(M_H^2)] \ . \tag{4.19}$$

Now eq. (4.19) and Table I imply that

$$\Delta r^H(\propto \alpha M_H^2) = \left(\frac{\alpha}{4\pi s^2} \xi \ln \xi\right) 2c^4 N \Delta(s^4 - s^2 - s^4 + s^2) = 0 \ . \tag{4.20}$$

Therefore there are <u>no</u> heavy Higgs corrections of order αM_H^2. This result has been noticed by many authors for the standard (one-doublet) model. It implies that radiative corrections will grow logarithmically rather than quadratically with the Higgs mass. Thus the effects of radiative corrections will be harder to see than one might naively expect.

V. THE GENERALIZED SCREENING THEOREM

A. Why There Are No $(\alpha M_H^2 \ln \xi)$ Corrections to M_W^2

In the last section we saw by explicit calculation that $\mathcal{O}(\alpha M_H^2 \ln \xi)$ corrections to M_W^2 cancelled. We will now elevate this observation to a screening theorem and discuss the general circumstances under which screening will occur.

As discussed in section III, the O(3) symmetry among the W^a's implies equal tree level masses $M^2_{W^1} = M^2_{W^2} = M^2_{W^3}$, and leads to the natural relation $M_Z/M_W = 1/c$. The screening theorem is a consequence of the equality of the three $\mathcal{O}(\alpha M_H^2 \ln \xi)$ radiative corrections: $\Delta M_{W^1}(M_H^2) = \Delta M_{W^2}(M_H^2) = \Delta M_{W^3}(M_H^2)$. Thus a redefinition of the parameter $\langle\phi\rangle \equiv v$ can be used to reabsorb the radiative corrections from vacuum polarization into the tree level Lagrangian.

The equality of the radiative corrections to the three W^a masses is evident from Table 1. Alternatively the equality can be seen by reexpressing the radiative corrections in terms of the product of two currents, where we divide the product into momentum dependent and momentum independent terms.

$$\langle J^a(p)J^a(-p)\rangle = \langle J^a(0)J^a(0)^+\rangle$$
$$+ [\langle J^a(p)J^a(-p)^+\rangle - \langle J^a(0)J^a(0)\rangle^+] \qquad (5.1)$$

Nothing in the momentum independent piece distinguishes the a = 1, 2 or 3 pieces and so this term preserves an $SU(2)_V$ symmetry. This is the term proportional to M_H^2. The second, momentum <u>dependent</u> term does distinguish between W^1 and W^3: W^1 has an on-shell external momentum of $p^2 = - M_W^2$ whereas W^3 has an on-shell external momentum of $p^2 = - M_Z^2$. This breaks the custodial $SU(2)_V$ symmetry.

Explicitly, the mass corrections are of the form:

$$\begin{aligned} \Delta M^2_{W^1} = \Delta M^2_{W^2} &= G(N)\ p^2\ \ell n\ \xi \Big|_{p^2 = M_W^2(\text{tree})} \\ \Delta M^2_{W^3} = \cdot\ &\quad G(N)\ p^2\ \ell n\ \xi \Big|_{p^2 = M_Z^2(\text{tree})} \end{aligned} \tag{5.2}$$

where G(N) includes coupling, combinatoric and group factors. Since $\Delta M_{W^1} \neq \Delta M_{W^3}$ the different mass corrections <u>cannot</u> be reabsorbed in the parameters of the tree level Lagrangian. For this reason $(\alpha M_W^2\ \ell n\ \xi)$ corrections to the tree level masses are physically significant and will be calculated in Section VI.

In general we can pose to any theory a "screen test:" do radiative corrections to the gauge boson masses have the same symmetry structure as the tree level masses? If so the theory has screening.

B. Illustration of the "Screen Test"

To illustrate how the screening theorem might fail to apply in another model, consider the field $\phi^T = (\phi_1, \phi_2, \phi_3)$ to be a single complex triplet. ϕ transforms as a (3,1) under $SU(2)_L \times U(1)_Y$ with hypercharge Y = - 1 and electric charge descending from 0 at the top of the multiplet to -2 at the bottom. This model does not have a custodial $SU(2)_V$ symmetry. We will find that the radiative corrections of $\mathcal{O}(\alpha M_H^2 \ln \xi)$ are not proportional to the tree level masses of the W^1, W^2 and W^3. Consequently the unscreened Higgs mass is physically significant.

Just as we did in Section II it is possible to define a C such that $CTC = -T^*$ and $C^2 = 1$ and such that $\tilde{\phi} \equiv C\phi^*$ transforms the same way as ϕ does under $SU(2)_L$. Using the two $SU(2)_L$ invariants $\tilde{\phi}^+\phi$ and $\phi^+\phi$ we can construct the most general $SU(2)_L \times U(1)_Y$ symmetry renormalizable, reflection invariant potential as:

$$U(\phi) = a(\phi^+\phi - f^2)^2 + b|\tilde{\phi}^+\phi|^2. \tag{5.3}$$

It follows that $\langle\phi_1\rangle = f$, $\langle\phi_3\rangle = \langle\phi_2\rangle = 0$ minimizes $U(\phi)$. Therefore,

$$\frac{M^2_{W^3}(\text{tree})}{M^2_{W^1}(\text{tree})} = \frac{(T^3)^2_{11}}{(T^1)^2_{11}} = \frac{1}{2} \tag{5.4}$$

The mass spectrum of the scalars in this theory is easily found to be

$$M^2_{\phi_3} = 4af^2 \text{ (real and imaginary parts of } \phi_3)$$

$$M^2_{\phi_1} = 4bf^2 \text{ (only real part of } \phi_1) \tag{5.5}$$

where we have shifted ϕ in the usual way.

Consider the simplifying limit of $b << a$ in which only the complex field ϕ_3 is massive. We therefore only need to calculate the one loop diagrams of Fig. 5.1. First we note that there are no $W^3\phi_3\phi_{1(2)}$ vertices because $(T^3)_{13} = (T^3)_{23} = 0$. Second, the two diagrams in Figs. 5.1a and b cancel with only ϕ_3 particles running around the loops (this is the same effect that prevents the π^+, for example, from renormalizing the photon mass). Therefore in this curious model $\Delta M_{W^3} = 0$ whereas $\Delta M_{W^1} \neq 0$. Hence,

$$\frac{\Delta M^2_{W_3}\text{ (1-loop)}}{\Delta M^2_{W_1}\text{ (1-loop)}} = 0 \neq \frac{M^2_{W_3}\text{ (tree)}}{M^2_{W_1}\text{ (tree)}} = \frac{1}{2} \tag{5.6}$$

Eq. (5.6) tells us that radiative corrections of $\mathcal{O}(\alpha M_H^2 \ln \xi)$ will be physical and thus that there is no screening.

VI. RADIATIVE CORRECTIONS TO M_W, M_Z OF ORDER $\alpha\, M_W^2\, \ell n\, \xi$

We have shown in section IV that 1-loop corrections to the gauge meson masses in the $U(N)_L \times SU(N)_R$ model are not proportional to $\alpha\, M_H^2\, \ell n\, \xi$. Instead the leading terms are proportional only to $\alpha\, M_W^2\, \ell n\, \xi$.

There are four such logarithmically enhanced contributions to Δr^H. One contribution which we label Δr^H(2-scalar) comes from the momentum dependent piece of loops with two scalars (see Fig. 6.1). The p^2 piece of these integrals was neglected in section IV because only the $\alpha\, M_H^2\, \ell n\, \xi$ contributions were of interest.

A second contribution to Δr which we label Δr^H(martian) comes from the notorious "martian" graphs (see Fig. 6.2). Next we must correct for the fact that in the calculation of Δr^H(2-scalar) all of the π^α fields were treated as if they were massless. This was an oversimplification -- three of them were really the three unphysical "eaten" Goldstone bosons which in 't Hooft-Feynman gauge should have been of mass M_W^2 (M_Z^2 for the neutral one). We thus must compute Δr^H(unphys) which arises from substituting a massive for a massless propagator in 3 of the 2-scalar graphs (see Fig. 6.3). Finally we compute $\pi^H_{\gamma\gamma}$, the correction to the photon propagator involving heavy scalars (see Fig. 6.5).

A. Momentum Dependent Part of Graphs with Two Internal Scalars

The most important contribution of $\mathcal{O}(\alpha\ M_W^2\ \ln\ \xi)$ to Δr^H is Δr^H(2-scalar) from graphs with two internal scalars. To calculate this quantity we return to consider the previously neglected terms proportional to $p^2 = -\ M_{W(Z)}^2$ in the graphs shown in Fig. 4.2.

The integrals proportional to p^2 in the standard integral $F(M_1M_2)$ are defined as $S'(M_1M_2)$ in Eq. (4.16-17). Note that from Eq. (4.16),

$$S'(P\pi)_{W(Z)} = S'(PP)_{W(Z)} = (-\ \tfrac{1}{3})\ (\tfrac{1}{4\pi})^2\ M_{W(Z)}^2\ \ln\ \xi_p. \quad (4.21)$$

Subscripts W(Z) denote corrections to the W(Z) propagators respectively. The Euclidean momentum p^2 is set equal to $-M_W^2$ and $-M_Z^2$ in the two cases.

For the group factors we can simply carry over the results we calculated in Table I, recalling that the tadpole graphs do not contribute because they are momentum-independent. When the 2-scalar group factors are multiplied by the standard momentum-dependent integral in eq. (4.16) we obtain the contributions to A_{WW} and A_{ZZ} that are displayed in Table II.

Feynman Diagrams of Type Shown in Fig. 6.1 with External W and The Following Virtual Particles	Contribution to A_{WW}
$\omega\pi$	$(G/3)\beta^2 c^2 [\Delta/N] M_W^2$
ηP	$(G/3)\beta^2 c^2 [\Delta/N] M_W^2$
$P\pi$	$(G/6)\beta^2 c^2 [N-4/N]\Delta M_W^2$
PP	$(G/12)\beta^2 c^2 (N\Delta) M_W^2$

Fig. 6.1 with External Z and the Following Virtual Particle	Contribution to A_{ZZ}
$\omega\pi$	$(G/3)\beta^2 [\Delta/N] M_Z^2$
ηP	$(G/3)\beta^2 [\Delta/N] M_Z^2$
$P\pi$	$(G/6)\beta^2 [N-4/N]\Delta M_Z^2$
PP	$(G/12)\beta^2 N\Delta(1+4s^4-4s^2) M_Z^2$

Table II. Two-Scalar Graphs.

$\Delta = \mathrm{tr}(T^3T^3)$, $\beta \equiv e/sc$, $G \equiv -(1/4\pi)^2 \ln \xi$.

Only non-zero contributions to A_{WW} and A_{ZZ} are listed.

It is then entirely straightforward to obtain Δr^H (2-scalar); we need only insert these contributions to A_{WW} and A_{ZZ} into eq. (4.4). This yields

$$\Delta r^H \text{ (2-scalar)} = \frac{N\Delta}{6} (3 - 2c^2) \frac{\alpha}{4\pi s^2} \ln \xi \quad . \tag{6.1}$$

Equation (6.1) corresponds to a choice of parameters such that all of the particles η, ω and P^α get a mass M_H. If on the other hand we choose our parameters such that only the ω particle acquires a mass M_H, then from Table II we need only retain the $\omega\pi$ terms. In this case eq. (4.4) reduces to:

$$\Delta r^H_\omega \text{ (2-scalar)} = \left(\frac{2\Delta}{3N}\right) \frac{\alpha}{4\pi s^2} \ln \xi \quad . \tag{6.2}$$

B. <u>Martian Graphs</u>

Martian graphs (see Fig. 6.2) come from terms in the Lagrangian with vertex factors of the form $WW\phi\langle\phi\rangle$. The relevant piece of the Lagrangian eq. (1.4) can be rewritten in terms of A at Z^0 as:

$$\mathscr{L}' = \mathrm{tr}\Big[\big(\partial\phi - i\beta(cT^1W^1\phi + cT^2W^2\phi + c^2T^3\phi Z^0 + s^2\phi T^3 Z^0$$

$$+ sc(T^3\phi - \phi T^3)A)\big) \times \text{(hermitian conj.)}\Big] \tag{6.3}$$

where we insert ϕ-v with $\langle\phi\rangle$ = v and ϕ defined in eq. (1.2); as before we define a standard integral that will appear in all of the graphs shown in Fig. 6.2:

$$G(M_{W(Z)}, M_H) \equiv \int \frac{d^4k}{(2\pi)^4} \left[\frac{1}{(p+k)^2} \frac{1}{(k^2+M_H^2)} \right]$$
$$= -\frac{1}{(4\pi)^2} \ln \xi . \tag{6.4}$$

As an example of the kind of coefficient we must determine, consider the vertex $W^1 v\phi Z^o$ coming from $tr\{\mathcal{D}\ddagger(\mathcal{D}\ddagger)^+\}$. Expanding the trace of the product we find that the term containing the W^1, Z^o and a vev is:

$$cv\beta^2 \; tr[\{T^3, T^1\}](\phi + \phi^+) \; W^1 Z^o. \tag{6.5}$$

It follows that W and Z^o couple only to the (massive) hermitian part of the ϕ field. Extracting the other vertices in a similar fashion the group identity

$$tr(CX^\alpha) \; tr(DX^\alpha) = tr(CD) - \frac{1}{N} \, trC \; trD \tag{4.12}$$

may be used to compute the coefficients of the graphs of Fig. 6.2. In Table III all of the Martian graph contributions to A_{WW} and A_{ZZ} are displayed. Using Table III we can sum all of the Martian graph contributions to A_{WW} and A_{ZZ}. When these are substituted into eq. (4.4) a little arithmetic yields

$$\Delta r^H(\text{Martian}) = \left(\frac{B-A}{\Delta}\right) \frac{\alpha}{4\pi s^2} \ln \xi . \tag{6.6}$$

Feynman Diagrams of Type Shown in Fig. 6.2 with External W^1 and The Following Virtual Particles	Contribution to A_{WW}
$W^1 P$	$2v^2\beta^4 c^4[B - 4/N\ \Delta^2]G$
$W^1 \omega$	$2v^2\beta^4 c^4[4/N\ \Delta^2]G$
$W^2 P$	$2v^2\beta^4 c^4\ A\,G$
ZP	$2v^2\beta^4 c^2\ A\,G$

Fig. 6.2 with External Z and The Following Virtual Particles	Contribution to A_{ZZ}
$W^1 P$	$2v^2\beta^4 c^2\ A\,G$
$W^2 P$	$2v^2\beta^4 c^2\ A\,G$
ZP	$2v^2\beta^4[B - 4/N\ \Delta^2]G$
$Z\omega$	$2v^2\beta^4[4/N\ \Delta^2]G$

Table III. Martian Graphs
$A \equiv tr[\{T^1,T^3\}]^2$, $B \equiv tr[\{T^3,T^3\}]^2$, $\Delta \equiv tr(T^3T^3)$,
$\beta \equiv e/sc$ and $G \equiv -(1/4\pi)^2 \ln \xi$.

If we maintain the larger symmetry by only giving the ω particle a mass eq. (6.6) reduces to

$$\Delta r_{\omega}^{H}(\text{Martian}) = \left(\frac{4\Delta}{N}\right) \frac{\alpha}{4\pi s^2} \ln \xi \quad . \tag{6.7}$$

In both cases we have used eq. (3.19) which tells us that $M_W^2(\text{tree}) = 2\beta^2 c^2 v^2 \Delta$.

The Feynman rules for the $AZ^0\phi$, $AW^1\phi$ and $AW^2\phi$ vertices are determined similarly from the Lagrangian, eq. (6.3). For example the $AW^{1(2)}\phi$ vertex is

$$\frac{sc^2\beta^2 v}{2} \operatorname{tr}\{T^{1(2)}[T^3,(\phi-\phi^+)]\} \tag{6.8}$$

which clearly picks out only the antihermitian part of ϕ. Since eq. (6.8) vanishes unless $[T^3,(\phi-\phi^+)]$ is proportional to $T^{1(2)}$, only that portion of $(\phi-\phi^+)$ proportional to $T^{2(1)}$ survives. But T^1 and T^2 are just the directions in the vacuum that correspond to two of the three eaten GB's. Consequently only the charged unphysical GB's $G^2(G^1)$ couple to the $W^{1(2)}$ and the photon simultaneously. We will see immediately below that in 't Hooft-Feynman gauge the unphysical Goldstone boson has a mass equal to that of the W. Therefore no large logs can occur in the photonic Martian diagrams (see Fig. 6.4). (The Martian with Z^0 external legs and a virtual photon can similarly be shown to couple only to the neutral unphysical Higgs.)

C. Loops Involving Unphysical Higgs

Thus far we have ignored the fact that three of GB's acquire a mass in 't Hooft-Feynman gauge. The mass arises as an artifact of the gauge fixing term $\mathscr{L}_{GF}$ in the Lagrangian.

$$\mathscr{L}_{GF} = \frac{1}{2\alpha}\,[\partial^\mu W^1 - G^a v\xi^{(1)}\,\mathrm{tr}(T^1X^a)]^2$$

$$+ \frac{1}{2\alpha}\,[\partial^\mu W^2 - G^a v\xi^{(1)}\,\mathrm{tr}(T^2X^a)]^2$$

$$\frac{1}{2\alpha}\,[\partial^\mu Z^o - G^a v\xi^{(2)}\,\mathrm{tr}(T^3X^a c^2 + X^aT^3 s^2)]^2 \qquad (6.9)$$

In 't Hooft-Feynman gauge $\alpha = 1$; $\xi^{(1)}$ is chosen equal to g and $\xi^{(2)}$ equal to β in order to cancel the mixed scalar-vector propagators. We take the $SU(2)_L$ generators T^a to be embedded in SU(N) such that

$$\mathrm{tr}(T^aX^\alpha) = \Delta^{1/2}\,\delta_{a\alpha} \qquad (6.10)$$

There is a simple way to take account of the unphysical Higgs. We only need to calculate the <u>difference</u> between the graphs calculated with massive unphysical Higgs, G^A, and with massless PGB's π^A (A = 1 - 3) that we erroneously used earlier. Let us call Δr^H(unphys) the correction due to the substitution of G^A for π^A. Only the momentum independent term from Sirlin's eq. (4.4) contributes:

$$\Delta r^H(\mathrm{unphys}) = -\frac{c^2}{s^2M_W^2}\,[c^2A_{ZZ} - A_{WW}], \qquad (6.11)$$

where A_{WW} (and A_{ZZ}) come from the graphs of Fig. 6.3. From eq. (6.9) the masses of the unphysical Higgs are:

$$M^2_{G^{1(2)}} = M^2_{W^{1(2)}} = 2v^2\beta^2c^2\Delta \text{ and} \tag{6.12}$$

$$M^2_{G^3} = M^2_{Z^o} = 2v^2\beta^2\Delta,$$

The momentum integral $E(M_1, M_2)$ for the graphs of Fig. 6.3 is therefore simply the difference between two S integrals (see eqs.4.16-17) that we have already calculated.

$$E(M_H, M_{W(Z)}) = S(M_H, M_{W(Z)}) - S(M_H, 0)$$

$$= \left(\frac{1}{4\pi}\right)^2 M^2_{W(Z)} \ln \xi + (\text{terms proportional to } M^2_H). \tag{6.13}$$

(The momentum-dependent piece is by definition included in S' and $S'(M_H, M_H) - S'(M_H, 0) = 0$; terms proportional to M_H^2 cancel by the screening theorem.)

Noting that the matrices X^α and T^A are normalized differently (eq. 6.10) we find that the vertex factors involving the unphysical Higgs are very similar to the vertices of Table III. We simply calculate the vertices using:

$$\phi = \frac{1}{\sqrt{2}}\, i\left((G^A - \pi^A)X^A + \ldots\right) \qquad A = 1\text{-}3 \tag{6.14}$$

where the elipses indicate the π^α fields $\alpha > 3$ and the P, ω and η fields. The contributions of all the graphs involving unphysical Higgs are shown in Table IV.

Feynman Diagrams of Type Shown in Fig. 6.3 with External W^1 and The Following Virtual Particles	Contribution to A_{WW}
$G^1 P$	$-(1/4\Delta)\beta^2 c^2 [B - 4\Delta^2/N] M_W^2\, G$
$G^1 \omega$	$-(1/4\Delta)\beta^2 c^2 [4\Delta^2/N] M_W^2\, G$
$G^2 P$	$-(1/4\Delta)\beta^2 c^2\, A\, M_W^2\, G$
$G^3 P$	$-(1/4\Delta)\beta^2 c^2\, A\, M_Z^2\, G$

Fig. 6.3 with External Z and The Following Virtual Particles	Contribution to A_{ZZ}
$G^1 P$	$-(1/4\Delta)\beta^2\, A\, M_W^2\, G$
$G^2 P$	$-(1/4\Delta)\beta^2\, A\, M_W^2\, G$
$G^3 P$	$-(1/4\Delta)\beta^2 [B - 4\Delta^2/N] M_Z^2\, G$
$G^3 \omega$	$-(1/4\Delta)\beta^2 [4\Delta^2/N] M_Z^2\, G$

Table IV. Graphs with Unphysical Higgs.
$\Delta \equiv \mathrm{tr}(T^3 T^3)$, $A \equiv \mathrm{tr}[\{T^1, T^3\}]^2$, $B = \mathrm{tr}[\{T^3, T^3\}]^2$, and $G \equiv -(1/4\pi)^2 \ln \xi$.

Once again it is straightforward to insert the unphysical Higgs' contribution to A_{WW} and A_{ZZ} into our formula (4.4) to obtain Δr^H(unphys):

$$\Delta r^H(\text{unphys}) = -\frac{(B-A)}{4}\frac{\alpha}{4\pi s^2}\ln\xi . \qquad (6.15)$$

For completeness we evaluate eq. (6.15) for the case in which only the ω is massive:

$$\Delta r^H_\omega(\text{unphys}) = -\left(\frac{\Delta}{N}\right)\frac{\alpha}{4\pi s^2}\ln\xi . \qquad (6.16)$$

D. Electric Charge Renormalization

In addition to the direct contributions to the W propagator, we must also renormalize the electric charge because α appears in the tree level gauge boson mass. Many contributions to the renormalization of e are contained in $\frac{\Delta r^{other}}{2}$ but the photon propagator can also be renormalized by heavy scalar loops of the type we have described here.

What pairs of particles could run around the loop in diagrams like Fig. 6.5? Any loops with neutral particles (ω and η) are excluded because they do not couple to the photon,A. We thus could have PP or Pπ. Expanding the Lagrangian in (eq. 6.3) reveals that the PπA vertex vanishes. This leaves only the PPA vertex made into a loop diagram in Fig. 6.5. The group factor is evaluated in the usual way, the standard integral is just $S'(M_P M_P)$, and we are left with:

$$\pi^{\mu\nu}_{\gamma\gamma} = -\left(\frac{N\Delta s^2}{3}\right)\frac{\alpha}{4\pi s^2}\ln\xi = \Delta r^H(\text{e-charge}). \qquad (6.17)$$

Recall that by definition $\pi^{\mu\nu}_{\gamma\gamma}$ is the coefficient of $(p^2 g_{\mu\nu} - p_\mu p_\nu)$.

E. No Infrared Catastrophe, No Ghosts

We will see in Section VII that among the N^2-1 PGB's are N-2 exact neutral GB's Bosons (EGB's), 1 neutral eaten GB, 2 charged eaten GB's and N^2-N-2 charged PGB's that get a small mass. One might worry that diagrams with two EGB's in a loop could give rise to a divergent momentum integration. However the $W\pi^{EGB}\,\pi^{EGB}$ vertex vanishes because the vertex is proportional to the commutator of two diagonal X^α generators.

One also might wonder about ghost loops of the type shown in Fig. 6.6. Since the ghosts have mass $\sim M_W$ there is no 1 TeV mass in the ghost loop graph in Fig. 6.6. Therefore the large logs of 2-scalar, martian and unphysical Higgs graphs dominate any contribution from ghost loops.

VII. OTHER TERMS IN THE CHIRAL LAGRANGIAN

Thus far we have not considered interactions other than those in the kinetic term, $\mathrm{tr}\{\mathscr{D}\phi(\mathscr{D}\phi)^{+}\}$. However the philosophy of effective field theories demands that we include all terms in which $SU(N)_L \times SU(N)_R$ chiral symmetry is broken by $SU(2)_L \times U(1)_Y$ gauge interactions. Using the non linear chiral representation we shall show in this section that terms other than those already considered in section VI are either suppressed by powers of α or else that they lack the logarithmic enhancement $\ell n\, \xi$.

As we saw in section II, when all the heavy fields are integrated out we are left only with the Goldstone Bosons, π^{α}. The exponential representation of the GB's is convenient in the generalized SU(N) case just as it was in the SU(2) case already considered in eq. (2.14):

$$\Sigma = \exp[i\pi^{\alpha}X^{\alpha}/f], \ \alpha = 1, \ldots, (N^2 - 1) \qquad (7.1)$$

f is of order Λ_{CSB} and the matrix Σ transforms under separate U(N) transformations L and R:

$$\Sigma \rightarrow L\,\Sigma\,R^{+}. \qquad (7.2)$$

If we had a dynamical TC theory we could discuss how gauge boson loops could renormalize the mass of techniquarks and hence of the PGB's. Such processes give rise to the famous vacuum alignment terms which specify the relative orientation of gauge generators and the chiral background generators [10,11].

In the language of effective field theory we can see the effect of gauge interactions on the PGB masses by looking for $SU(N)_L \times SU(N)_R$ breaking effective mass terms in the chiral Lagrangian formalism. For example we might try to construct a term using two left-handed gauge generators:

$$f^4 g^2 \; \mathrm{tr}\{\Sigma T_L^a \Sigma^+ \; \Sigma T_L^a \Sigma^+\} = f^4 g^2 \; \mathrm{tr}(T_L^a T_L^a) \tag{7.3}$$

Eq. (7.3) is chiral $SU(N)_L \times SU(N)_R$ invariant and so cannot give mass to the W directly. It is also independent of Σ and so cannot give mass to the π's and thereby <u>indirectly</u> contribute to M_W. For the same reason two hypercharge (T_Y) generators do not contribute either. Suppose we take one left generator and one hypercharge generator:

$$f^4 g g' \; \mathrm{tr}(\Sigma T_L^a \Sigma^+ T_Y). \tag{7.4}$$

This term is forbidden under the symmetries of the Lagrangian $g \to -g$ and $g' \to -g'$. Therefore no PGB masses are generated in order $g^2 \Lambda_{CSB}^2 = \alpha \Lambda_{CSB}^2 = M_W^2$.

However, we can construct a term without derivatives using four $SU(2)_L \times U(1)_Y$ generators:

$$A\, g^2\, g'^2\, \mathrm{tr}(\Sigma T_L^a \Sigma^+ T_Y\, \Sigma T^a \Sigma^+ T_Y). \tag{7.5}$$

(A has dimension Λ_{CSB}^4 because the term is a vacuum energy density.) Therefore a small mass of order $M_{PGB}^2 = \alpha^2 \Lambda_{CSB}^2 = \alpha M_W^2$ is given to the PGB's. Other terms also contribute to the charged PGB masses in this order. It is easily shown that all of the charged PGB's get a mass squared of $M_{PGB}^2 \sim \alpha M_W^2$.

It is interesting to note that not all of the GB's get mass -- the neutral scalars i.e. those in the vacuum symmetric directions

$$T^3,\ (T^3)^2,\ (T^3)^3,\ \ldots,\ (T^3)^{N-1} \tag{7.6}$$

are exactly massless. One can see this by considering a subspace of the vacuum orientations

$$\Sigma^k_{exact} \equiv \exp\left[i\, \frac{\pi^k (T^3)^k}{f}\right], \tag{7.7}$$

where k is a superscript on Σ^k_{exact} and π^k and an exponent on T^3. Now recall that $\Sigma T_Y = -\,\Sigma T_3$. Thus when we insert Σ^k_{exact} into terms like eq. (7.3) we can commute Σ^k_{exact} through T_Y and find

$$f^4 g^2 g'^2 \ tr(T_L^a T_R T_L^a T_R) = (7.3) \tag{7.8}$$

which has no field dependence.

Other $SU(2)_L \times U(1)_Y$ terms that break $SU(N)_L \times SU(N)_R$ involve more factors of T_L^a and T_Y which carry additional powers of α. In this sense the terms are small compared to those calculated in section VI.

The small $(\alpha M_W^2)^{1/2}$ PGB masses contribute to M_W^2 through radiative corrections that are only of order $\alpha^2 M_W^2 \ \ln \alpha$. Graphically we indicate the effective mass-generating interaction by the blob which we insert in a scalar loop with two π fields as shown in Fig. 7.1. Fortunately we have already calculated a graph with two massive scalars of this type in Section IV.

To switch to the non linear representation we simply use Λ_{CSB}^2 as the cut-off for the internal momentum integration. This gives us (suppressing Feynman parameters) a coefficient of $g_{\mu\nu}$ of the form:

$$A_{WW}(p^2) \sim \alpha \ (p^2 + M_{PGB}^2) \ \ln \left[\frac{p^2 + M_{PGB}^2}{\Lambda_{CSB}^2} \right] . \tag{7.9}$$

p^2, the external momentum, is set equal in M_W^2. Since

$$M_{PGB}^2 \sim \alpha M_W^2 << p^2 = M_W^2$$

we can simply neglect the small masses acquired by the PGB's in eq. (7.9). Similarly, we can neglect the PGB mass in $A_{ZZ}(p^2)$ because it has the same basic form as $A_{WW}(p^2)$ except we set the external momentum $p^2 = - M_Z^2$. Finally to find $A_{WW}(p^2 = 0)$ we simply set $p^2 = 0$ in eq. (7.9):

$$A_{WW}(p^2=0) = \alpha M_{PGB}^2 \ \ell n \ \frac{M_{PGB}^2}{M_Z^2} \sim \alpha^2 M_W^2 \ \ell n \ \xi \qquad (7.10)$$

which is smaller than the contributions of section VI by a factor of α. Since by eq. (4.4) Δr^H depends only on $A_{WW}(M_W^2)$, $A_{ZZ}(M_Z^2)$ and $A_{WW}(0)$ we conclude that the PGB masses in this model are too small by a factor of α to figure in our calculation of the W and Z masses. All other effective mass terms for the π's have more gauge generators and therefore are suppressed by even more powers of α.

In general what happens if we add more covariant derivatives? Lorentz invariance requires that any term have an even number of covariant derivatives. By the definition of $\mathcal{D}$ either ∂_μ or igTW acts on Σ. (Here let the gauge group sum implicity over $SU(2)_L$ and $U(1)_Y$.) If ∂_μ is applied to Σ one of the two factors is extracted: an internal momentum k_μ/f or an external momenta p_μ/f. The f arises from the definition of Σ (eq. 7.1). Thus any term with additional ∂'s extracting external momenta will be suppressed by factors of $p^2/f^2 = M_W^2/\Lambda_{CSB}^2 \sim \alpha$. If an internal momentum is extracted then the momentum integral will no longer give a large logarithm. Instead we would get $\int_0^{\Lambda_{CSB}} d^4k \, \frac{(k^{2m})}{k^4}$ for m > 1.

Let us now examine more carefully the order α graphs that are log enhanced in order to pursue the connection between the linear and non linear pictures. Recall $M_H^2 \sim \Lambda_{CSB}^2 \sim 1\ \text{TeV}^2$. In Section VI we explicitly calculated the graphs in Figs. 7.2. In the non linear representation all of the heavy scalars are contracted to a point. Thus we have the corresponding diagrams (of Fig. 7.3).

The point is this: in the linear case the large log comes from the ratio of the heavy scalar (M_H^2) to the renormalization scale ($\mu^2 = M_W^2$). In the non linear case the large log comes from the ratio of the external momentum (M_W^2) to the cutoff at the chiral symmetry breaking scale (Λ_{CSB}^2). Note the complementarity between the two representations: the small graphs in one representation are the large contributions in the other.

Similarly, the Martian graphs exhibit the same complementarity as one passes from the linear to the non linear representation. The linear graphs are shown in Fig. 7.4.

Contracting the heavy scalar in Fig. 7.4a to a point we get the non-linear version of the same diagrams. (See Fig. 7.5).

In the linear representation the graph with a heavy scalar (Fig. 7.5a) gets its logarithmic enhancement as before; the graph is momentum independent and its αM_W^2 coefficient comes from the extra factor of $g^2v^2 = M_W^2$. The linear representation of the Martian graph with a π field (see Fig. 7.4b) has no large log enhancement because it has no propagator with a large mass.

By contrast, in the non-linear picture the Martian graph with a heavy scalar line contracted appears as a pure 4-gauge particle interaction (Fig. 7.5a) and is suppressed by α. The non-linear representation also includes a Martian graph with a π field (Fig. 7.5b). On dimensional grounds the term corresponding to this graph has a Λ_{CSB}^2 coefficient. It is also logarithmically enhanced by the ratio of the Λ_{CSB} cutoff to the external momentum squared, $p^2 = -M_W^2$.

We could phrase the connection between the two representations another way. When, in the linear representation, we renormalize at $\mu^2 = \Lambda_{CSB}^2$ we get the same contributions to Δr^H diagram by diagram as we get in the non linear representation cut off at Λ_{CSB}. Clearly Δr^H must be independent of our choice of μ in a given representation. When we set $\mu^2 = M_W^2$ in the linear representation some terms get large logs that had none before; others that were large become small. The sum of terms composing Δr^H remains, however, unchanged.

In sum we have seen in Section VI how to compute certain corrections to the W and Z^o masses. We have now shown that these are indeed the only $\mathcal{O}(\alpha)$ logarithmically enhanced terms in the effective theory.

VIII. CONCLUSION

There are two ways to phrase the conclusion. Suppose the $p\bar{p}$ experiment at CERN discovers that the $W^{\pm}$ and Z^{0} masses are within a few percent of their predicted electroweak values in a theory without heavy scalars. Then we could rule out strongly coupled Higgs representations with dimension N greater than some small N_{max}. In the event that the gauge bosons appear significantly above 82.0 GeV (see eq. (4.8)), one could then invoke heavy scalars with large weak isospin to account for the large W mass. In either case it is useful to express our results in the form of a functional dependence of M_W on N.

In the renormalization scheme used in this paper, the physical mass of the W is given by,

$$M_W \text{ (physical)} = \sqrt{\frac{\pi\alpha}{\sqrt{2}\, G_\mu s^2}} \left[1 + \frac{\Delta r^{other}}{2} + \frac{\Delta r^{H}}{2}\right] \qquad (8.1)$$

where $s^2 = 0.229$, $\alpha = 137^{-1}$, $G_\mu = 1.17 \times 10^{-5}\ \text{GeV}^{-2}$ and $\frac{\Delta r^{other}}{2} = 0.051$. $\Delta r^{other}/2$ represents corrections from leptons, hadrons and that part of weakly interactions bosons of ghosts that are independent of $\frac{M_H^2}{M_W^2} = \xi$. $\frac{\Delta r^H}{2}$ represents the contribution of the heavy scalars.

Let us assume that all of the heavy scalars have approximately the same ($\sim$1 TeV) mass. We can collect the contributions to Δr^H in eqs. (6.1), (6.6), (6.15), and (6.17):

$$\begin{aligned}\frac{\Delta r^H}{2} &= \frac{1}{2}\,[\Delta r^H(\text{2-scalar}) + \Delta r^H(\text{Martian}) \\ &\quad + \Delta r^H(\text{unphys}) + \Delta r^H(\text{e-charge})] \\ &= \frac{1}{2}\,[\frac{N\Delta}{6} + \frac{3}{4}\,(\frac{B-A}{\Delta})]\,\frac{\alpha}{4\pi s^2}\,\ln\,\xi \\ &= (0.007N^4 + 0.143N^2 - 0.225)\;0.012,\end{aligned} \tag{8.2}$$

where we have used $\frac{B-A}{\Delta} = \frac{2N^2-3}{5}$, $\Delta = \frac{N(N^2-1)}{12}$ and $\xi = M_H^2/M_W^2 = \alpha^{-1} = 137$. As a check we can consider the special case of two doublets (N=2). We can then use the chiral representation to write ϕ as

$$\phi = \frac{1}{\sqrt{2}}\,[\omega + iT^a\pi^a + i(\eta + iT^bP^b)] \tag{8.3}$$

Only the ω particle gets a vev. Since we have given the η a mass degenerate with the P^b's, the second doublet does not even break $SU(2)_L \times U(1)_Y$. It therefore does not contribute the W mass through radiative corrections. Consequently we expect the model to reproduce the result Sirlin [8] eq. (4.6) derived for a doublet with a single real heavy field. Indeed putting N=2 in eq. (8.2) yields:

$$\frac{\Delta r^H}{2}\ (N=2) = (\tfrac{11}{24})\ \frac{\alpha}{4\pi s^2}\ \ln\ \xi. \tag{8.4}$$

Finally we can also check that in the limit where the second doublet is massless we recover the same result. This is the case when we choose the coefficients in the scalar potential such that the ω particle is the only massive particle. In this case: we have from eqs. (6.2), (6.7), and (6.16):

$$\begin{aligned}\frac{\Delta r^H_\omega}{2} &= \left(\frac{\Delta}{3N} + \frac{2\Delta}{N} - \frac{\Delta}{2N}\right)\ \frac{\alpha}{4\pi s^2}\ \ln\ \xi \\ &= \left[\frac{11(N^2-1)}{72}\right]\ \frac{\alpha}{4\pi s^2}\ \ln\ \xi.\end{aligned} \tag{8.5}$$

For N=2 this gives

$$\frac{\Delta r_\omega^H}{2}\ (N=2) \quad = \quad \left(\frac{11}{24}\right) \frac{\alpha}{4\pi s^2}\ \ell n\ \xi = 0.006 \tag{8.6}$$

which again reproduces, as it should, the standard single doublet model with a heavy Higgs.

The CERN $p\bar{p}$ collider group [12] has recently reported a measurement of the W mass to be

$$M_W^{CERN} = 81 \pm 4 \text{ GeV} . \tag{8.7}$$

If we restrict $r^H/2$ to predict a M_W^{theor} within three standard deviations of M_W^{CERN} then, by eqs. (4.3) and (4.8):

$$82 + (77.9)\ \Delta r^H/2 \leq 93 \tag{8.8}$$

and so,

$$\frac{\Delta r^H}{2} \leq 0.141 . \tag{8.9}$$

From Table V we see that eq. (8.9) implies that

$$N \leq N_{max} = 5 . \tag{8.10}$$

N	$\Delta r^H/2$	$\Delta r^H_\omega/2$
2	0.006	0.006
3	0.020	0.015
4	0.046	0.028
5	0.093	0.044
6	0.167	0.064
7	0.282	0.088
8	0.449	0.116
9	0.683	0.147

Table V

Suppose for some reason that only the ω particle gets the 1 TeV mass. Then the restriction that M_W^{theor} be within three standard deviations of M_W^{CERN} implies

$$\Delta r^H_\omega/2 \leq 0.141. \tag{8.11}$$

Consulting Table 4 we see that

$$N_\omega \leq N_{\omega\ max} = 9. \tag{8.12}$$

In conclusion, the natural relation $M_Z/M_W = 1/\cos\theta_W$ suggested by neutral current experiments, together with a direct measurement of M_W, severely restricts any model with an effective $SU(N)_L \times SU(N)_R$ Higgs symmetry (e.g. technicolor models) which breaks $SU(2)_L \times U(1)_Y$ in the simple manner described here.

Figure Captions

4.1 Self-energy diagram calculated in a Euclidean metric.

4.2 Diagrams that separately contribute to the W mass in order $\alpha M_H^2 \ln \alpha$ for $M_H^2/M_Z^2 \sim \alpha^{-1}$. Together, all such contributions cancel.

4.3 $W^1 P^\alpha P^\beta$ vertex.

5.1 Diagrams leading to radiative corrections to the W mass of $\mathcal{O}(\alpha M_H^2 \ln \xi)$ in a toy model with a single charged isotriplet.

6.1 2-scalar graph which produces momentum dependent contribution to the W mass that is of order $\alpha M_W^2 \ln \xi$.

6.2 Martian graphs in which the antennae designate vevs.

6.3 Unphysical Higgs graphs in which G^A are the massive unphysical Higgs: $M^2_{G^{1(2)}} = M_W^2$; $M^2_{G^0} = M_Z^2$. $G^A - \pi^A$ designates the difference between two propagators: one with unphysical Higgs mass equal to M_W^2 (or M_Z^2) and one with massless π^A field.

6.4 Martian graphs with virtual photons have only a virtual gauge boson and a virtual unphysical Higgs with mass M_W. Therefore in the linear representation there is no log ξ enhancement.

6.5 Electric charge renormalization. Only the $P^\alpha P^\beta A$ vertex includes a heavy scalar and the photon. All other such vertices vanish.

6.6 Ghost loop.

7.1 Typical contribution of PGB masses to radiative correction of gauge boson mass.

7.2 2-scalar contributions to W mass in linear representation. M_H^2 designates any of the heavy fields ω, P or η; the π fields are massless.

7.3 Non-linear representation of 2-scalar graphs. Only π-π loop contributes.

7.4 Linear representation of Martian graphs. Only (a) with heavy scalar contributes with log ξ.

7.5 Non-linear representation of Martian graphs. Only (b) contributes in order α.

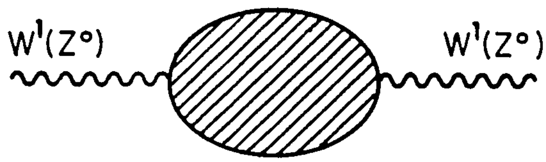


Fig. 4.1

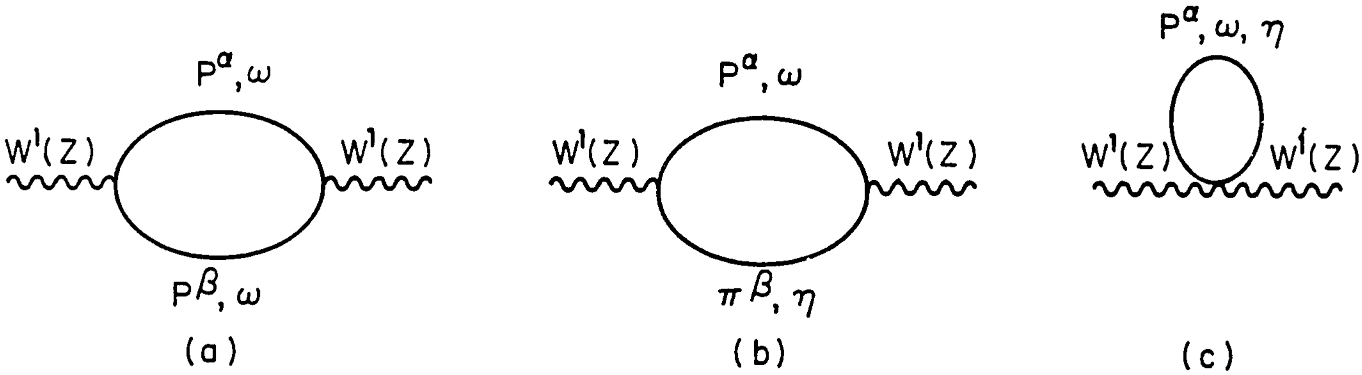


Fig. 4.2

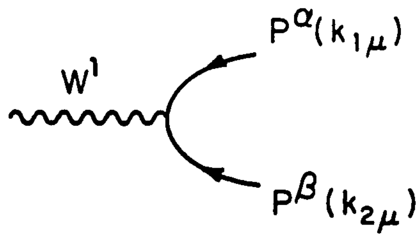


Fig. 4.3

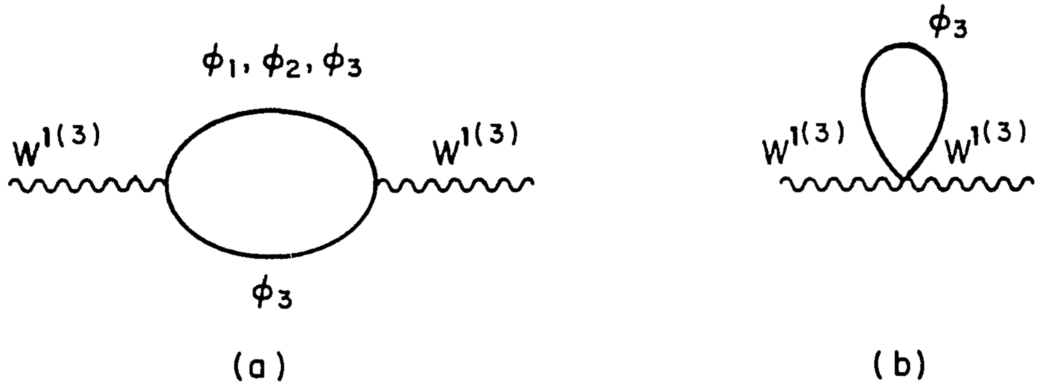


Fig. 5.1

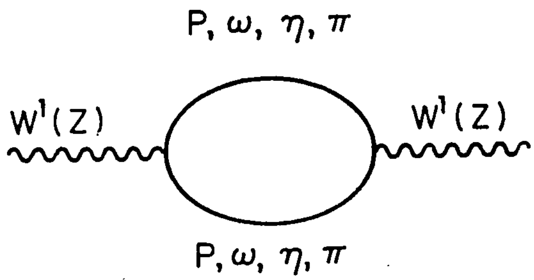


Fig. 6.1

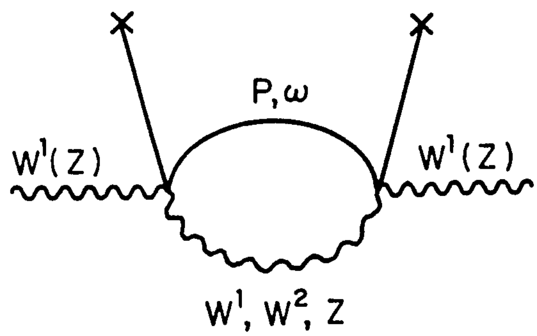


Fig. 6.2

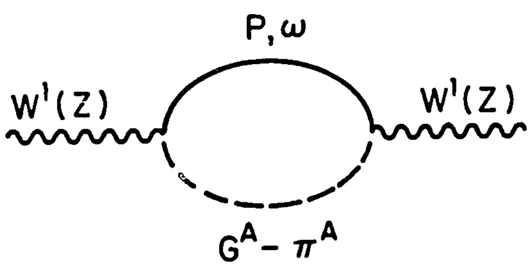


Fig. 6.3

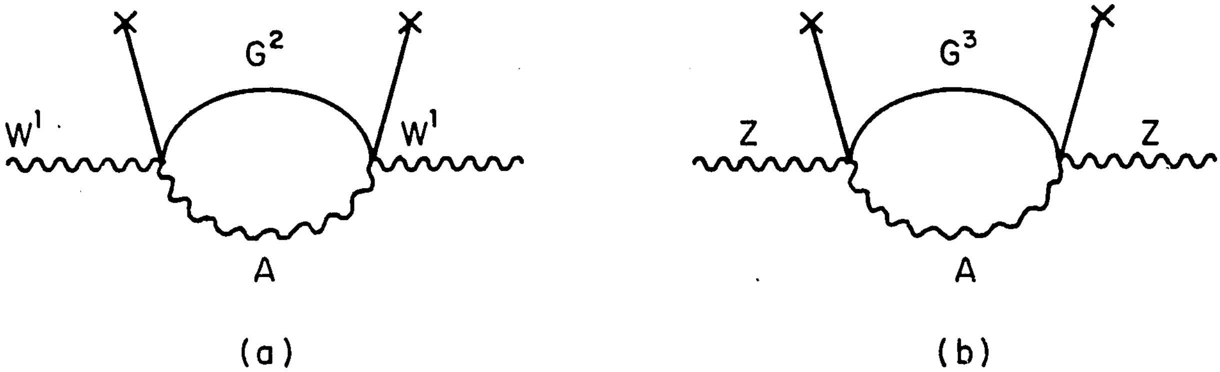


Fig. 6.4

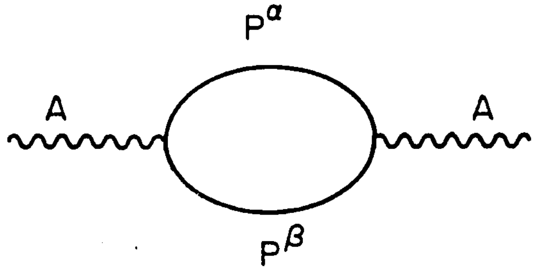
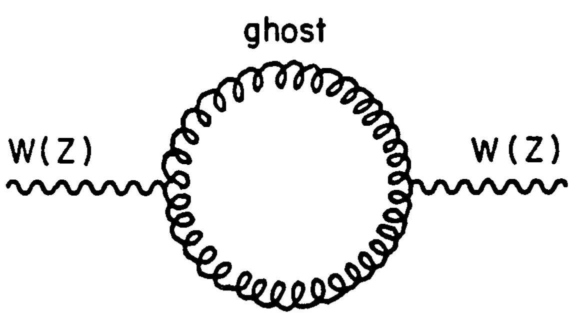


Fig. 6.5



Fig, 6.6

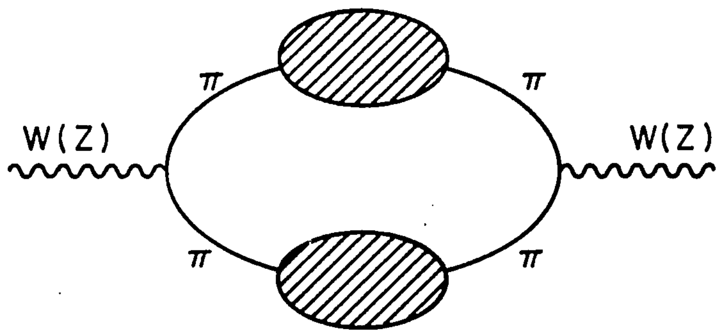


Fig. 7.1

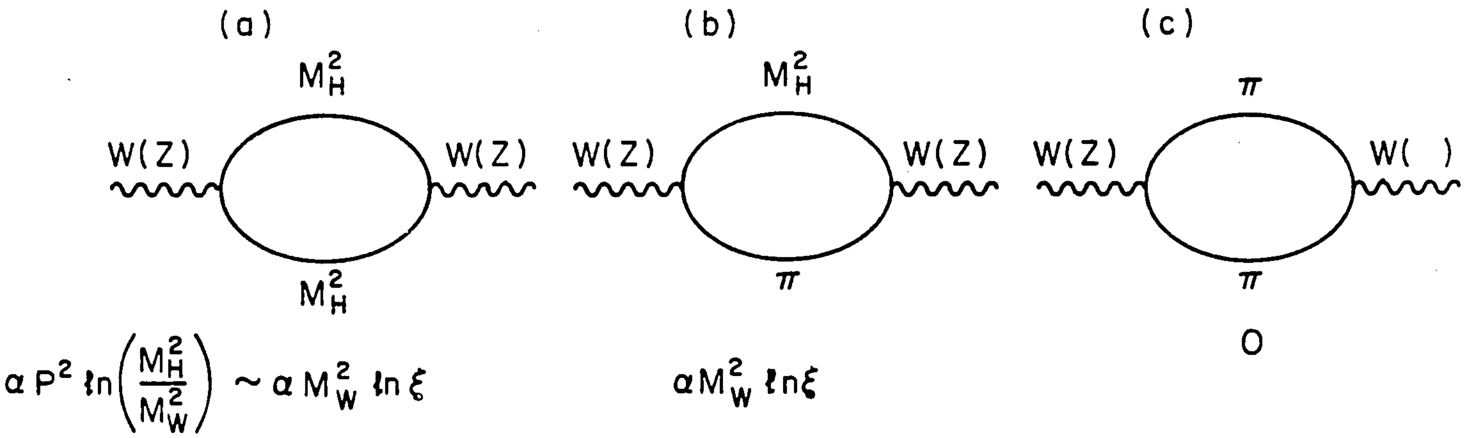


Fig. 7.2

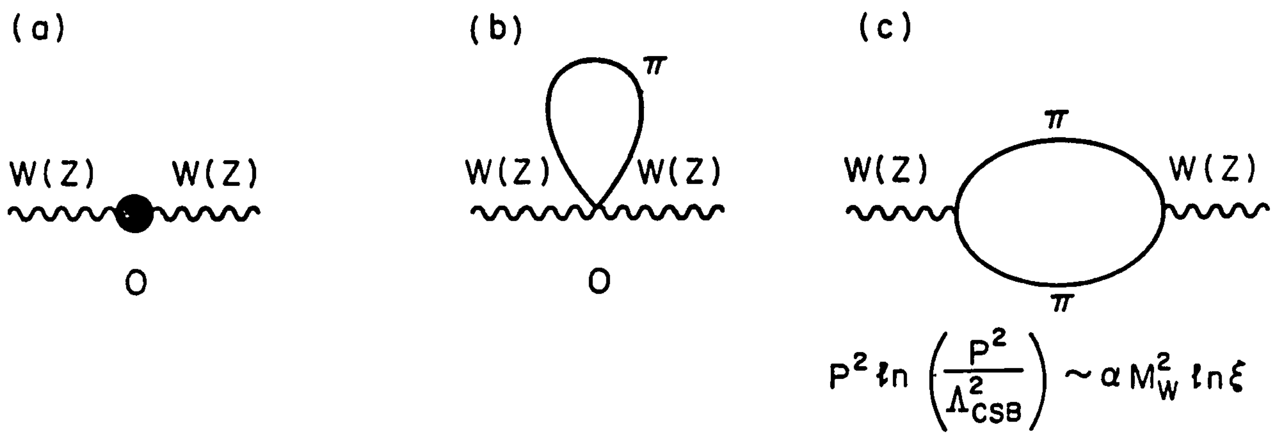
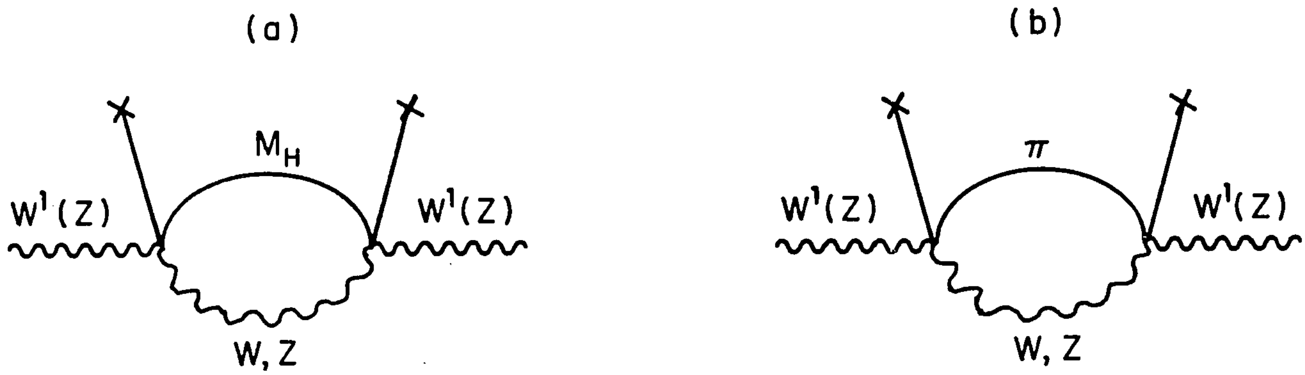
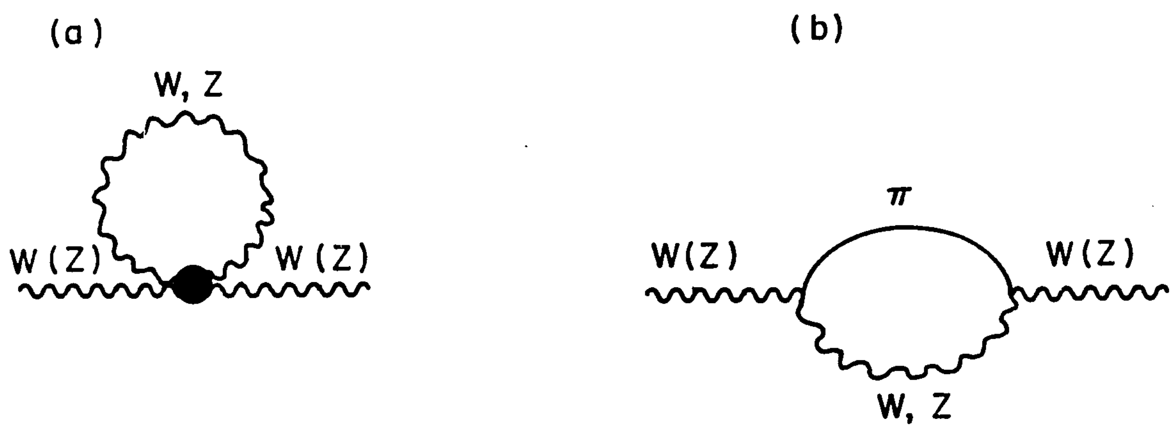


Fig. 7.3



$$\alpha g^2 V^2 \ln\left(\frac{M_H^2}{M_W^2}\right) \sim \alpha M_W^2 \ln \xi \qquad \alpha M_W^2 \ln\left(\frac{M_W^2}{M_W^2}\right) \sim 0$$

Fig. 7.4



$$\alpha^2 M_W^2 \ln \xi \sim 0 \qquad \alpha^2 \Lambda_{CSB}^2 \ln\left(\frac{M_W^2}{\Lambda_{CSB}^2}\right) \sim \alpha M_W^2 \ln \xi$$

Fig. 7.5

APPENDIX

A-1) PGB MASSES

In section VII we saw that it was possible to construct effective mass terms for the π fields that were of order α^2. (There were none of order α.) In fact there are two such terms, one of which was discussed in eq. (7.5):

$$f^4 g^2 g'^2 \; [\mathrm{tr}(\Sigma T_L^a \; \Sigma^+ T_Y)]^2 \tag{A.1}$$

and

$$f^4 g^2 g'^2 \; \mathrm{tr}(\Sigma T_L^a \; \Sigma^+ T_Y)^2 \tag{A.2}$$

The f's are inserted on dimensional grounds where $f^2 = \Lambda^2_{CSB} = 1\ \mathrm{TeV}^2$ as before. We wish to show that these terms give mass to all and only the charged PGB's, i.e. the neutral GB's are exact.

Expanding eq. (A.1) we get

$$f^4 g^2 g'^2 \; [\mathrm{tr}(T^a\{T^3 + i[\underset{\sim}{\pi}, T^3] + \frac{i^2}{2} [\underset{\sim}{\pi}, [\underset{\sim}{\pi}, T^3]] + \ldots\}]^2,$$

with

$$\underset{\sim}{\pi} = \frac{\pi^a X^a}{f} \quad a = 1 \text{ to } N^2\text{-}1. \tag{A.3}$$

keeping only terms quadratic in π results in,

$$V_1(\text{quad}) \equiv - f^4 g^2 g'^2 \{[\mathrm{tr}(T^a[\underset{\sim}{\pi}, T^3])]^2$$

$$+ \mathrm{tr}(T^a T^3)\ \mathrm{tr}(T^a[\underset{\sim}{\pi}, [\underset{\sim}{\pi}, T^3]])\}. \qquad (A.4)$$

Similarly the terms quadratic in π from eq. (A.2) are given by

$$V_2(\text{quad}) \equiv - f^4 g^2 g'^2 \{\mathrm{tr}(T^a[\underset{\sim}{\pi}, T^3])^2$$

$$+ \mathrm{tr}(T^a T^3 T^a\ [\underset{\sim}{\pi}, [\underset{\sim}{\pi}, T^3]]). \qquad (A.5)$$

The π masses can now be determined by evaluating the four traces in (A.4) and (A.5). We choose the SU(N) generators in the defining representation. Let E_{ij} be the matrix with components a,b:

$$(E_{ij})_{ab} \equiv \frac{1}{\sqrt{2}} \delta_{ai}\ \delta_{jb}. \qquad (A.6)$$

The Cartan subalgebra of SU(N) is generated by the matrices H_M where

$$(H_m)_{ij} \equiv \left(\sum_{k=1}^{m} \delta_{ik}\ \delta_{jk} - m\ \delta_{i,m+1}\ \delta_{j,m+1}\right) / \sqrt{2m(m+1)}$$

In this appendix we normalize the SU(N) generators according to

$$\mathrm{tr}(X^\alpha X^\beta) = \frac{1}{2}\,\delta_{\alpha\beta} \tag{A.8}$$

where an X^α is either one of the E_{ij}'s or one of the H_m's. The SU(2) generators T^a in the spin-p representation ($2p + 1 = N$, where N is the dimension of multiplet) are:

$$(T^3)_{ij} = (p - j + 1)\,\delta_{ij} \equiv D_j\,\delta_{ij}$$

$$(T^-_{ij}) = (T^+)_{ij} = \sqrt{j(n-j)}\,\delta_{i,j-1} \equiv c_j\delta_{i,j-1} \tag{A.9}$$

Thus

$$\mathrm{tr}(T^3T^3) = N(N^2-1)/12 \equiv \Delta$$

and so,

$$\mathrm{tr}(X^\alpha T^a) = \sqrt{\Delta}\,\delta_{\alpha a}. \tag{A.10}$$

The traces of (A.3) and (A.4) can now be evaluated. The first trace in (A.3) is

$$f^4g^2g'^2\,[\mathrm{tr}(T^a[\underset{\sim}{\pi},\,T^3])]^2 \tag{A.11}$$

$$= f^4g^2g'^2\,\mathrm{tr}(\underset{\sim}{\pi}[T^3,\,T^a])\,\mathrm{tr}(\underset{\sim}{\pi}[T^3,\,T^a]) \tag{A.12}$$

$$= 2f^2 g^2 g'^2 \, tr(X^\alpha T^+) \, (X^\beta T^-) \, \pi^\alpha \pi^\beta \qquad (A.13)$$

$$= 2\alpha M_W^2 \, tr(T^+ \, E_{i+1,i}) \, tr(T^- \, E_{j,j+1}) \, \pi_{i+1,i} \, \pi_{j,j+1} . \qquad (A.14)$$

$$= \alpha M_W^2 \, c_i c_j \, \pi_{i+1,i} \, \pi_{j,j+1} \qquad (A.15)$$

where in going from (A.13) to (A.14) we assumed the relation $\alpha f^2 \sim M_W^2$, consistent with the text.

Similarly, the second trace in eq. (A.3) is

$$\alpha M_W^2 \, tr(T^a T^3) \, tr(T^a [\underset{\sim}{\pi}, \, [\underset{\sim}{\pi}, \, T^3]])$$

$$= \alpha M_W^2 \Delta \, tr(T^3 [E_{ij} [T^3, \, E_{k\ell}]]) \, \pi_{ij} \, \pi_{k\ell}$$

$$= \alpha M_W^2 \Delta \, (D_\ell - D_k) \, tr(T^3 [E_{ij}, \, E_{k\ell}]) \, \pi_{ij} \, \pi_{k\ell}$$

$$= \alpha M_W^2 \Delta \, (D_\ell - D_k)(D_i - D_j) \, tr(E_{k\ell} \, E_{ij}) \, \pi_{ij} \, \pi_k \qquad (A.16)$$

$$= \frac{\alpha M_W^2 \, \Delta}{2} \, (\ell - k)^2 \, \pi_{k\ell} (\pi_{k\ell})^T . \qquad (A.17)$$

The first trace of eq. (A.5) is

$$\alpha M_W^2 \, tr(T^a [\underset{\sim}{\pi}, \, T^3])^2$$

$$= \alpha M_W^2 \, tr(T^a [T^3, \, E_{ij}] \, T^a [T^3, \, E_{k\ell}]) \, \pi_{ij} \, \pi_{k\ell} \qquad (A.18)$$

$$= \alpha M_W^2 \; (D_i - D_j)(D_k - D_\ell) \; tr(T^a E_{ij} \; T^a E_{k\ell}) \; \pi_{ij} \; \pi_{k\ell} \tag{A.19}$$

For a=3 (A.18) becomes

$$\alpha M_W^2 \; (D_i - D_j)(D_k - D_\ell) \; D_i D_k \; tr(E_{ij} E_{k\ell}) \; \pi_{ij} \; \pi_{k\ell} \tag{A.20}$$

$$= \frac{\alpha M_W^2}{2} \; (k-\ell)^2 \; D_k D_\ell \; \pi_{k\ell} (\pi_{k\ell})^T . \tag{A.21}$$

for a = +, - (A.18) gives

$$2\alpha M_W^2 \; (D_i - D_j)(D_k - D_\ell) \; tr(T^+ E_{ij} \; T^- E_{k\ell}) \; \pi_{ij} \; \pi_{k\ell}$$

$$= 2\alpha M_W^2 \; (j-1)(\ell-k) \; c_{i-1} c_k \; tr(E_{i-1,j} \; E_{k+1,\ell}) \; \pi_{ij} \; \pi_{k\ell} \tag{A.22}$$

$$= \alpha M_W^2 \sum_{k,\ell} \; (k-\ell)^2 \; c_k c_\ell \; \pi_{\ell+1,k+1} \; (\pi_{\ell,k})^T \tag{A.23}$$

Together the T^+,T^- and T^3 pieces yield:

$$\sum_{k\ell} \alpha M_W^2 \; (k-\ell)^2 \; \{D_k D_\ell \; \pi_{k\ell} (\pi_{k\ell})^T$$

$$+ \; c_k c_\ell \; \pi_{\ell+1,k+1} (\pi_{\ell k})^T \} . \tag{A.24}$$

Finally, the second trace of eq. (A.5) is

$$\alpha M_W^2 \; tr(T^a T^3 T^a [\underset{\sim}{\pi}, \; [\underset{\sim}{\pi}, \; T^3]]) . \tag{A.25}$$

Note that

$$T^aT^3T^a = [T^a,\ T^3]\ T^a + T^3T^aT^a = (1+\Delta)T^3 \qquad \text{(A.26)}$$

so (A.25) is just equal to

$$(1+\Delta)\,\mathrm{tr}(T^3[\underset{\sim}{\pi},\ [\underset{\sim}{\pi},\ T^3]]) \qquad \text{(A.27)}$$

where the trace in (A.27) is just that computed in (A.16).

To summarize: the first contribution to the PGB mass matrix, (A.11), gives mass to the charge-1 PGB's; the rest, (A.16), (A.18) and (A.25) give mass to all of the charged PGB's. Thus all of the charged PGB's acquire a mass of order αM_W^2 as advertised.

References

[1] H. Georgi and I. N. McArthur, Nucl. Phys. B202 (1982) 382.

[2] H. Georgi and S. L. Glashow, Phys. Rev. Lett. 47 (1981) 1511.

[3] A. C. Longhitano, Phys. Rev. D22 (1981) 1166.

[4] A. C. Longhitano, Nucl. Phys. B188 (1981) 118.

[5] R. Renken and M. E. Peskin, Nucl. Phys. B211 (1983) 93.

[6] H.-S. Tsao, AIP Conference Proc. No. 72 Subseries No. 23. *Weak Interactions as Probes of Unification*, ed. C. B. Collins, L. N. Chang and J. R. Ficenec (Virginia Polytechnic Institute, 1980).

[7] M. Veltman, Acta Phys. Polon. B8 (1977) 475.

[8] A. Sirlin, Phys. Rev. D22 (1980) 971.

[9] W. J. Marciano and A. Sirlin, Cornell Z^0 Workshop, February 6-8, 1981, CLNS 81-485, pp. 40-60.

[10] J. Preskill, Nucl. Phys. B177 (1981) 21.

[11] M. E. Peskin, Nucl. Phys. B175 (1980) 197.

[12] G. Arnison *et al.*, CERN preprint EP/83-13 (1983).

ACKNOWLEDGEMENTS

I would most of all like to thank Howard Georgi and John Preskill for advice and encouragement during my several years in the Physics Department. Their contagious enjoyment of the subject rescued me many times from discouragement. It has been a real pleasure to have worked with them.

I am also grateful to Steven Weinberg who has been supportive of my work throughout my graduate studies. Specifically, several of our discussions led to my original interest in models with large weak isospin.

My fellow Junior Fellows have been an unending source of help, abuse and amusement, especially Mark Wise, Luis Alvarez-Gaume and Paul Ginsparg. I would also like to thank the other theoretical graduate students, above all Aneesh Manohar, Mark Claudson, Ann Nelson, Steven Della Pietra, David Kaplan, and Phil Nelson for friendship and a nearly infinite number of useful discussions.

The Ad Hoc Committee on Physics and History of Science (H. Georgi, E. Heibert, G. Holton, and J. Preskill) have with wisdom and sympathy helped me through the most complicated course of study that I could invent.

I would like to thank the Society of Fellows for support and the opportunity to explore a new discipline outside the one in which I was originally trained.

My thanks are due also to Carol Davis for her excellent typing.

Finally I would like to thank my family and friends for putting up with me, two theses and more Higgs than they could possibly have ordered.