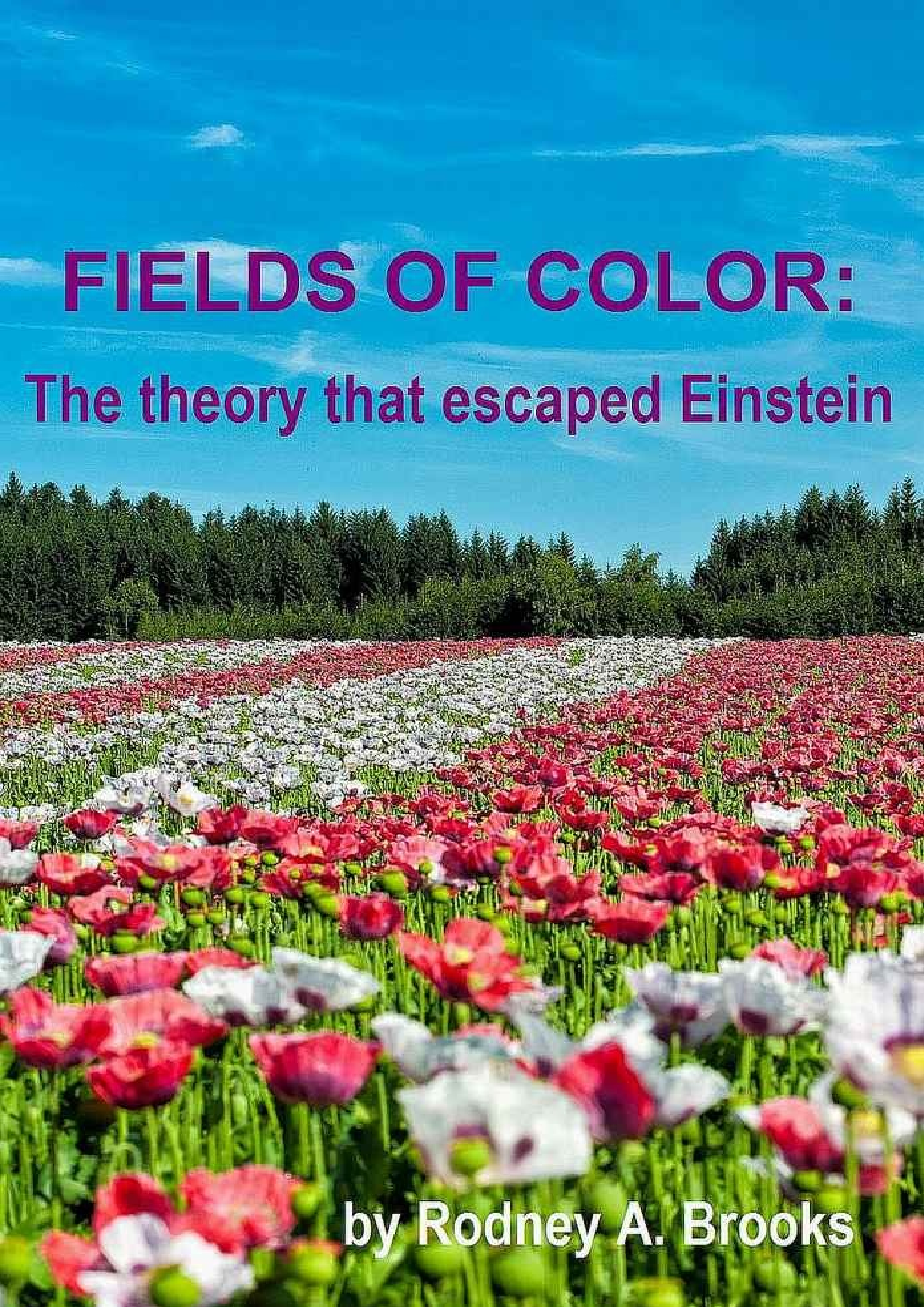


FIELDS OF COLOR:

The theory that escaped Einstein

by Rodney A. Brooks

FIELDS OF COLOR:
The theory that escaped Einstein

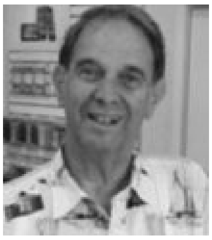
By Rodney A. Brooks

2nd edition 2011
3rd edition 2016

Published by Rodney A. Brooks

Printed by Universal Printing, LLC Silver Spring, Maryland

ISBN 978-0-473-17976-2



Rodney A. Brooks was born in Syracuse, NY, in 1932. He attended the University of Florida and Harvard University, where he received his Ph.D. in physics with Nobel laureate Norman Ramsey in 1963. During 25 years at the National Institutes of Health in Bethesda, MD, he published 124 refereed articles, including the design of a high-resolution PET scanner and the invention of dual-energy computed tomography. As an amateur clarinetist he founded and led a klezmer band called Shir Delite. After retiring in 1999 he and his wife lived for nine years in New Zealand, where he began work on this book.

DEDICATION

I dedicate this book to three people who played a special part in its development. First is my wife Karen, who not only supported the project throughout, but also wrote the first paragraph and in Chapter 7 outdid Einstein (see “Karen vs. Albert”). Second is Prof. Edward Finn, the godfather of the book, who urged me to write it, gave it its title (now its subtitle), and provided invaluable help in finishing it. But most of all, I dedicate this book to the memory of Julian Schwinger, one of the greatest physicists of all time and, sadly, one of the most forgotten. It was Schwinger who turned Quantum Field Theory into the beautiful structure that I have tried to convey to a wider public.

TABLE OF CONTENTS

PREFACE

> **People like you and me, though mortal, of course, like everyone else, do not grow old no matter how long we live. What I mean is that we never cease to stand like curious children before the great Mystery into which we are born. – Albert Einstein (E1979, p. 82)**

When I was 14, I read Arthur Eddington's The Nature of the Physical World. I remember how excited I was to learn that a desk is not a solid piece of furniture. If you could look closely enough, said Eddington, beyond the power of any microscope, you would see that it is made of atoms so tiny that we can hardly grasp their tinyness. These atoms are made of even tinier nuclei, and these in turn are made of particles called neutrons and protons, while little electrons whiz around them in orbits like miniature solar systems. So that is what the world is made of, I thought, and this was the first step in my eventual decision to make physics my career. Of course it didn't occur to me then to wonder what electrons, protons and neutrons are made of.

Some time later, as a graduate student, I attended a three-year lecture series at Harvard University by Julian Schwinger. The timing was perfect. Schwinger's development of Quantum Field Theory (QFT) had matured and he was about to publish a monumental work, "A theory of the fundamental interactions". I sat mesmerized, as did others.

> **Attending one of [Schwinger's] formal lectures was comparable to hearing a new major concert by a very great composer, flawlessly performed by the composer himself… The delivery was magisterial, even, carefully worded, irresistible like a mighty river… Crowds of students and more senior people from both Harvard and MIT attended… I felt privileged – and not a little daunted – to witness physics being made by one of its greatest masters. – Walter Kohn, Nobel laureate (M2000, p. 593-4)**

As Schwinger stood at the blackboard, writing ambidextrously and speaking mellifluously in well-formed sentences, it was as if God Himself was handing down the Ten Commandments. The equations were so elegant that it seemed the world couldn't be built any other way. From the barest of principles, he derived the equations of QFT, even including the gravitational field. Not only was the mathematics elegant, but the philosophic concept of a world made of *properties of space* seemed to me much more satisfying than Eddington's mysterious particles. I was amazed and delighted to see how the paradoxes of relativity theory and quantum mechanics that I had found so baffling disappeared or were resolved.

Later on, I must admit, things got more complicated as the number and variety of fundamental fields grew, and quarks entered the scene. But to my knowledge, QFT remains as the true fabric of which the world is made. What's more, I believe it is the only fabric of which the world *could* be made.

Unfortunately, Schwinger, once called "the heir-apparent to Einstein's mantle" by J. Robert Oppenheimer, never had the impact he should have had on the world of physics or on the public at large. Instead, the more colorful and outgoing Richard Feynman came to the fore. It is his image, not Schwinger's, that is enshrined on a postage stamp. It is possible that Schwinger's very elegance was his undoing.

> **Julian Schwinger was one of the most important and influential scientists of the twentieth century… Yet even among physicists, recognition of his fundamental contributions remains limited, in part because his dense formal style ultimately proved less accessible than Feynman's more intuitive approach. However, the structure of modern theoretical physics would be inconceivable without Schwinger's manifold insights. His work underlies much of modern physics, the source of which is often unknown even to the practitioners. His legacy lives on not only through his work, but also through his many students, who include leaders in physics and other fields. – J. Mehra and K.A. Milton (M2000, p. v)**

In the 50 years that passed since my student days, I have seen very little mention of QFT in its true fields-only sense. Instead I have seen a bombardment of books and articles that keep repeating the paradoxes that people are expected to accept. Physical intuition has disappeared or, worse yet, is sneered at. Far from bringing to the public an understanding of nature, these popular books and articles have brought confusion and chaos. This hit me hard one day as I was reading Joseph Heller's memoir, Now and Then. Heller is the author of Catch 22, one of my all-time favorites, and when I read that *he* tried to understand quantum mechanics and had to give it up (see quote in Chapter 1), I knew that something was badly wrong. And so I decided to write a book.

My mission soon turned into a labor of love, with emphasis on labor. I had not anticipated the breadth and depth of the subject, or the drama as our greatest minds waged what I think is our greatest battle: to understand the world we find ourselves in (a far more worthy battle than the wars we are so good at waging against each other). By drama, I mean not only the philosophic struggle to wrest nature's secrets from their most hidden recesses with only the flimsiest of evidence. I also mean the human side of the story — stories that are sometimes tragic, sometimes nettlesome, but always fascinating.

This book is my attempt to bring to the public the same satisfaction and understanding that I felt in Schwinger's courses, and to dispel the paradoxes of physics that prevent so many people from understanding the natural world. The book is aimed at two audiences. The primary audience consists of those who, like Heller, have made some attempt to understand modern physics and found it to be incomprehensible. The other audience is those people who have not read much about physics, but who would like to learn about it in a way they can understand. I hope that my efforts will bear fruit and that the reader will come away from the book feeling that nature is not mysterious or paradoxical, but is understandable and indeed makes perfect sense.

Rodney Brooks
Wanaka, New Zealand, 2010

PREFACE TO 2ND EDITION

The second edition brings Fields of Color up-to-date with a description of more recent developments in the *standard model*, otherwise known as *Quantum Chromodynamics*. Of particular importance is the calculation of hadron masses from basic field equations, which has been called “one of the greatest scientific achievements of all time” – an achievement of which the public is largely unaware! A discussion of dark matter is also included, along with other clarifications and improvements.

Rodney Brooks
Prescott, Arizona, 2011

PREFACE TO 3RD EDITION

The third edition is a major rewriting. Among the new features are an explanation of gravity waves and their detection, and solutions to the measurement problem (Schrödinger's cat), the Ehrenfest paradox, and the mystery of dark matter.

Rodney Brooks
Silver Spring, Maryland, 2016
rodney@brooks2.com

CHAPTER 1
INTRODUCTION

I'm interested in learning things. I hope I can finally understand physics before I leave the earth. – Bill Clinton (MSNBC interview, Feb. 17, 2011)

Welcome to the world of physics — a world you may once have thought you would never enter. Perhaps you were told that present-day physics is too hard and so you didn't try. Or perhaps you read a popular article or book and found the subject to be incomprehensible. Well, I don't blame you; you are not the only one. The way modern physics is usually presented, it *is* incomprehensible. Despite the plethora of popular books on science, there are few people who can do more than mouth the words they have read.

And yet there is a theory that makes perfect sense and that can be understood by anyone. I am referring to Quantum Field Theory (QFT), a theory that has been overlooked or ignored in most books about physics. This theory, in its true sense of "no particles, only fields," can return us to the good old days when every educated person had a conceptual grasp of science, without needing any mathematical ability.

<u>The good old days</u>. In 1666, when Isaac Newton discovered his law of gravity, people did not need to understand the calculus that Newton invented so that he could calculate the orbit of the moon. It was enough to know that there is a universal force that causes apples to fall to the ground and that this same force holds the moon in its orbit around the earth. There was nothing paradoxical about this concept; indeed it explained so many things that the poet Alexander Pope wrote as Newton's epitaph:

Nature and Nature's laws lay hid in night.
God said, "Let Newton be!" And all was light.

Two centuries later, James Clerk Maxwell developed equations that united electric and magnetic forces into a single field called (naturally) the electromagnetic (EM) field, and he showed that light is a travelling oscillation in this field. These concepts also could be grasped and understood by anyone, even those who could not solve Maxwell's equations.

The discovery of atoms at the beginning of the 20th century posed a more difficult challenge. As Sir Arthur Eddington wrote, after referring to his "ordinary" writing table as Table No. 1:

> **Table No. 2 is my scientific table. It is a more recent acquaintance and I do not feel so familiar with it. It does not belong to the world previously mentioned – that world which spontaneously appears around me when I open my eyes… It is part of a world which in more devious ways has forced itself on my attention. My scientific table is mostly emptiness. Sparsely scattered in that emptiness are numerous electric charges rushing about with great speed; but their combined bulk amounts to less than a billionth of the bulk of the table itself. – A. Eddington (The Nature of the Physical World, 1927, p. ix)**

Yet there was nothing in this new picture that is paradoxical or that violates common sense, and it didn't take long for people to understand and accept atomic theory. This is an example of how, when evidence demands it, our intuitive picture of things can be replaced by a less intuitive picture that we accept *because* of common sense.

EINSTEIN’S ENIGMAS

This pleasant state of affairs ended when Albert Einstein entered the scene. Einstein was without doubt the great light of 20th century physics. His contributions are immense and he is seen as the embodiment of ultimate wisdom.[1] But it is also true that Einstein’s contributions caused much confusion:

> **…for all of his popular appeal and surface accessibility, Einstein also came to symbolize the perception that modern physics was something that ordinary laymen could not comprehend, “the province of priest-like experts”. — W. Isaacson (I2007, p. 6)**

The situation was so bad that poet J.C. Squire added the following lines to Pope’s couplet (A2004, p. 253):

It did not last: the Devil, howling “Ho!
Let Einstein be,” restored the status quo.

The reasons for the confusion – this curtain of chaos cast over physics — are three theories for which Einstein was either wholly or partly responsible: Special Relativity,[2] General Relativity, and Quantum Mechanics (QM). Of course Einstein did not develop the theory of QM — in fact, he eventually repudiated it. However, in his 1905 Nobel-prize paper he introduced its central concept: *wave-particle duality*. While Einstein may not be the father of QM, he surely is its grandfather.

I refer to these theories as “enigmas” because they involve concepts that just don’t make sense to most people — concepts that are truly paradoxical. To testify to the seriousness of the problem, I call the following witnesses to the stand: Marilyn vos Savant, Chaim Weizmann, and Joseph Heller, with guest appearances by Elsa Einstein, Margot Einstein and Richard Feynman.

Special Relativity. Marilyn vos Savant has the highest IQ score ever recorded, according to the Guinness Book of World Records. The following question and answer appeared in her column “Ask Marilyn”:

> **Q. Is there any way for the average person to grasp the theory of relativity? A. In my opinion, no… Without being highly educated in physics, we can only read summaries of the theory, accept the**

> **points on faith and then successfully repeat what we've learned to others. — M. Vos Savant (Parade magazine, Sept. 9, 2001)**

Einstein's wife Elsa also had trouble with the theory. When asked if she understood relativity, she replied simply, "Understanding relativity is not necessary for my happiness" (I2007, p. 246). If neither Einstein's wife nor the smartest person in the world could understand relativity, something must surely be wrong. The problem is the many paradoxical statements that arise from the theory. For example, take the statement that nothing can go faster than light. A sensible person can't help but ask, "Why on earth not?"

General Relativity. Chaim Weizmann, a brilliant chemist and the future president of Israel, once accompanied Einstein on a trans-Atlantic sea voyage. After the voyage was over, Weizmann reported:

> **Einstein explained his theory to me every day, and on my arrival I was fully convinced that he understood it. – C. Weizmann (S1986, p. xi)**

Nor did Einstein have any better luck with his step-daughter, Margot Einstein. While working on a book for the lay reader about relativity,

> **…he read every page out loud to Elsa's daughter Margot, pausing frequently to ask whether she indeed got it. "Yes, Albert," she invariably replied, even though (as she confided to others) she found the whole thing totally baffling. – W. Isaacson (I2007, p. 232)**

The problem with general relativity, as it is usually presented, is that gravity is attributed to *curvature* in four-dimensional *space-time*. What is this thing called space-time, one asks, and how can it be curved? Curvature of a two-dimensional surface can be pictured, but curvature in three dimensions, let alone four, is beyond most people's grasp.

Quantum Mechanics. Joseph Heller was one of America's great 20th century novelists — author of the classic Catch 22. Having written brilliantly about the psychological world, he felt that as an educated, intelligent human being he should also know something about the physical world. His attempt, however, went sadly wrong.

> **In recent years I have been… writhing in an exasperating quandary over quantum mechanics, which, to my mind, remains impossible even to define, let alone comprehend. - J. Heller (Now and Then, p. 194)**

He was not alone. Richard Feynman, a leading 20th century physicist, said:

> **I think I can safely say that nobody understands quantum mechanics. — R. Feynman (The Character of Physical Law, 1965)**

According to QM, matter is made of particles that do things no self-respecting particle could ever do — things that only waves can do. The idea that particles can behave like waves is a concept that even physicists find paradoxical:

> **[Wave-particle duality] is a paradox because particles are, by definition, localized entities that follow definite trajectories while waves are not confined to any particular path or region of space. How could the same thing be both confined and not confined, both a particle and a wave? – T. Norsen ("Intelligent design in the classroom", Forum on Physics and Society, vol. 35, no. 3, 2006)**

According to a physicists' joke, "light is waves on Mondays, Wednesdays and Fridays; it is particles on Tuesdays, Thursdays and Saturdays; and on Sundays we think about it." The situation is so bad that philosophers invented a philosophy called *logical positivism*, which states essentially that hope must be abandoned of ever describing a consistent reality.

> **I am a positivist who believes that physical theories are just mathematical models we construct, and that it is meaningless to ask if they correspond to reality, just whether they predict observations. – S. Hawking (P1997, p.169)**

Einstein's search. Even as the developers of QM proceeded further with particles and wave-particle duality, Einstein came to believe that reality must consist of fields and only fields. He spent the last 25 years of his life searching for a unified field theory — a theory that would not only combine gravity and EM forces (the only two forces known at the time) into a single field, but would also include matter.

> **What appears certain to me, however, is that in the foundations of any consistent field theory, there shall not be, in addition to the concept of field, any concept concerning particles. The whole theory must be based solely on partial differential equations and their… solutions. – A. Einstein (The theory of relativity, and other essays, p. 34)**

Quite naturally, he took as his starting point the gravitational equations that had brought him so much fame and acclaim, trying to expand them in a way that would include the EM field and also matter. He tried first one mathematical approach, then another. At one point his friend Wolfgang Pauli admonished him,

> **“Within a year, if not before, you will have abandoned that whole *distant parallelism,* just as earlier you gave up the *affine theory*…” The following January, [Einstein] admitted to Pauli, “So you were right after all, you rascal.” – W. Isaacson (I2007, p. 344)**

Einstein’s search was not successful.

> **And so it went for another two decades… To the very end, he struggled to find his elusive unified field theory. And the final thing he wrote, before he went to sleep for the last time, was one more line of symbols and numbers that he hoped might get him, and the rest of us, just a little step closer to the spirit manifest in the laws of the universe. – W. Isaacson (I2007, p. 344, 543)**

THE SOLUTION

While Einstein was pursuing his lonely and unsuccessful search for a unified field theory, other physicists were taking a different track that led eventually to QFT. This wonderful theory, so little known or appreciated, fulfills Einstein's desire for a world made only of fields, except that forces and matter are not combined into a single field, as Einstein hoped. In fact, there are at least seven different types of field. Not only is QFT the answer to Einstein's search, it also resolves his enigmas, and in a way that can be understood by the man (or woman) on the street (see Chapters 7-9).

What is a field? Abandoning the familiar picture of solid particles and replacing it with intangible fields is not easy. It will require a leap of imagination greater than the atomic picture that Eddington struggled with. To put it briefly, a field is a *property* or a *condition* of space. The field concept was introduced into physics in 1845 by Michael Faraday as an explanation for electric and magnetic forces. His experiment with iron filings that align themselves in the region around a magnet is done today by every physics student (Fig. 1-1).

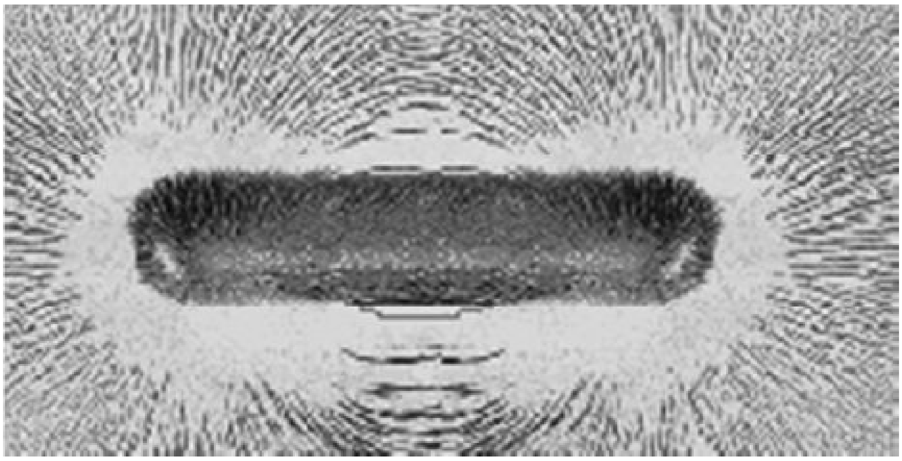


Fig. 1-1. Iron filings forming "lines of force" about a magnet led Faraday to believe that there must be a field present in the space around the magnet (www.ndt-ed.org).

However, the idea that fields can exist by themselves as properties of space was too much for physicists of the time to accept. Instead, they invented an invisible substance called *ether* to carry the EM oscillations. Belief in the ether prevailed for decades, but when no evidence for its existence could be found, despite many attempts, the ether was finally abandoned and physicists accepted that the EM field has an existence in itself. The idea that space can have properties does not come easily, but by the time you finish this book you will be comfortable with the concept of fields.

Field equations. In 1864 James Maxwell developed equations that describe how the strength of the EM field changes with time. These changes are *local* in the sense that only field intensities at the point of interest affect what happens at that point. However, a change of field strength somewhere else can propagate through space and eventually reach the point of interest, just as a stone dropped in water creates waves that propagate through the water. Light travels through space the same way, with a change in the EM field at one point creating changes at adjacent points, and so on. The way that fields propagate and the speed at which they propagate are determined by the field equations.

What is a quantum? In the centennial year of 1900 Max Planck introduced the idea that the EM field is not a continuous "classical" field, but is made of pieces, or chunks, that he dubbed *quanta* (from Latin *quantum* meaning "how much"). While Maxwell's classical EM field can be arbitrarily small, quantum fields are made of chunks that cannot be further reduced. Quanta may overlap each other, but each one maintains its separate identity; it lives a life and dies a death of its own. In that sense, and in that sense only, field quanta resemble particles.

"Particles" are quanta. In the 1920's it was found that the particles that make up matter exhibit wave characteristics. This led to the development of QM, with its characteristic wave-particle duality. This nettlesome problem (see Norsen quote above) was solved when QFT came along. In QFT there are no particles; there are only fields. It was Julian Schwinger who, in 1954, completed the formulation of QFT by finally treating force fields and matter fields on an equal basis.

> **…these two distinct classical concepts [particles and fields] are merged and become transcended in something that has no classical counterpart – the quantized field that is a new conception of its own, a unity that replaces the classical duality. – J. Schwinger (S2001)**

QFT vs. QM. It is important to understand that QFT is not just a variation of QM. Conceptually, the two couldn't be more different. QM describes a world made of particles and the equations give the probability that a particle is at a given point. QFT describes a world made of fields and the equations give the strength of the field at a given point. Another difference is that QM equations deal with numbers, while the fields of QFT are described by vectors in an abstract space known as *Hilbert space*. Fortunately, it is not necessary to understand the equations or to understand Hilbert space in order to grasp the concepts of QFT.

Fields of color. A special feature of this book is the use of color to depict fields that in themselves are unpicturable. Just as the color blue permeates the sky, so you will be asked to picture space as permeated with color, with different colors representing different fields. Color is an appropriate tool for this job because color is something that does not exist in itself; it exists only as a property of something else. By using colors to represent physical fields, we remind ourselves that fields are a property *of* space, not a separate substance *in* space.

The choice of color, of course, is arbitrary. I chose blue to represent the gravitational field, so the gravitational field of the earth is visualized as a blueness in the surrounding space — a blueness that extends not just to the sky, but to the moon and beyond, becoming fainter at greater distances. Similarly, the EM field is pictured as a greenness of space, while the two nuclear force fields — strong and weak — are represented by purple and brown. Finally, I chose yellow and red for the two matter fields, so “particles” like the electron and proton are pictured not as small spheres but as smeared-out blobs of yellow and red.

Limitation of color. There is a limit to how far the use of color can take us. A single color can help us visualize a field that is otherwise invisible, but it cannot inform us about the complexity of the field — and the fields we will encounter are complex. The EM field, for example, contains both electric and magnetic components, each of which has a direction associated with it, yet we will picture it simply as a green field. Also the use of color cannot convey the quantum nature of the fields – the fact that each field is made up of many individual pieces that keep their own identity, overlapping as they may be. The colors can help us picture the fields as “being there”, but they don’t tell the whole story.

THE BATTLE

When the theories of relativity were experimentally confirmed, the red carpet was rolled out despite (or perhaps because of) the fact that almost no one understood them. Einstein received world-wide acclaim unlike that bestowed on any other scientist, and was named “Man of the Century” by Time Magazine. QM, on the other hand, crept into public awareness without fanfare, but then gathered momentum. Today it, more than relativity, dominates articles about science, epitomizing the incomprehensibility of modern physics (see Candorville comic strip on p. 129). Now QFT represents as big a revolution in our thinking as relativity and QM put together — in fact, it combines and explains them. Yet you will look long and hard to find a book for the public that says quantized fields are the fabric of the cosmos.

A battle of three-rounds. There have been three rounds in the battle between fields and particles, and the battle is still going on. The first round came in 1905 when Einstein reintroduced Newton’s concept of light as corpuscles to explain the photoelectric effect, for which he received the Nobel prize (Chapter 3). The second round took place in the 1920’s when Schrödinger’s hope for a field theory of matter was abandoned because of the particle-like behavior that physicists could not ignore (Chapter 6). The third round occurred in the 1940’s when Feynman’s particle approach to quantum electrodynamics triumphed over Schwinger’s field picture, partly because Feynman’s method was easier to use (Chapter 6).

Einstein rejects QFT. Feynman was not the only physicist to reject QFT. Albert Einstein also rejected it after asking a friend, Valentin Bargmann, to give him a private tutorial on the subject. This is not surprising; at the time QFT was neither complete nor successful, nor was Bargmann its best spokesman. We must also consider Einstein’s great mistrust of the QM theory that preceded QFT — a mistrust based on its probabilistic nature. The idea of randomness in physics was repugnant to Einstein, as expressed in his famous comment, “I cannot believe that God plays dice with the universe.” As his friend and biographer, Abraham Pais, wrote:

> **I know from experience how difficult it was to discuss quantum field theory with him. He did not believe that… quantum mechanics provided a secure enough basis for relativistic generalizations. Relativistic quantum field theory was repugnant to**

> **him… and he did not believe in any of its consequences. – A. Pais (<u>P1982</u>, p. 463)**

<u>What if?</u> I can't help but wonder what might have happened if Einstein had waited a bit longer for his tutorial and chosen a different tutor – perhaps Julian Schwinger.

> **The year is 1954, the year Schwinger published the final installment of his "Theory of Quantized Fields" papers. The 36-year-old Schwinger comes to Princeton to meet the 75-year-old Einstein, who is to die the next year. "Prof. Einstein," he begins, in his calm but self-assured manner, "I want to show you a theory that unites all the known forces, as well as matter, into a field theory that incorporates both the principle of relativity as you have formulated it and the quantum principle that you helped introduce. It is a theory that is philosophically satisfying and, more importantly, has produced agreement with experiment to an undreamed of precision." "Well, go ahead then," says Einstein, and Schwinger proceeds to expound his theory in Einstein's Princeton study.**
>
> **The sessions are intensive, with Einstein asking many probing questions along the way. Finally Einstein says, "This is indeed a beautiful theory, but it seems there are six separate fields – four force fields and two matter fields. [The Higgs field was not known at the time.] As you know, my hope was to find a single field that comprises all forces and matter. This theory of yours does not meet that objective." "True," says Schwinger, "but may I remind you what your friend Pauli once said: 'What God hath put asunder, let no man join together.' If the God you often refer to chose to have six fields, it is not for us to second-guess Him. Yes, there are six different fields, each with its own properties and equations, but this is the way the universe seems to be made. As Niels Bohr once said to you, "It is not for us to tell God what to do."**
>
> **"All right," says Einstein, "I can accept the six fields, but there is a further problem that is more troubling to me, and that is the probabilistic nature of your theory. I have never been able to believe that God would play dice with the universe." "I understand," says Schwinger, "and I admit that QFT has not solved the causality problem. But at least it has localized the**

probability aspect to a single event called quantum collapse – an event that is not covered by the equations of QFT. That is to say [Schwinger's favorite expression], QFT itself is causal, but it only takes us so far. At some point a discontinuous event takes place that is not covered by the theory. It is possible that at some time a theory will be developed for quantum collapse and it may turn out to be causal after all. On the other hand, it may be that it is a truly probabilistic process, and perhaps God does play dice with the universe. Either way, it doesn't detract from what QFT has accomplished."

Einstein gives in. He is captivated by Schwinger's elegance and the soundness, both conceptual and mathematical, of his theory. He recognizes that QFT meets his desires for a field theory without "putting together what God hath torn asunder". He appreciates the explanation that QFT provides for his two theories of relativity. He then announces to the world that QFT is the theory he has been looking for all his life. Of course this turns the tide and QFT becomes accepted throughout the physics and public communities, restoring the popular appreciation and understanding of science that once existed in the time of Newton and Maxwell.

But that didn't happen.

And so the most wonderful theory ever devised — a theory that has produced more precise agreement with experiment than anything before, that encompasses all forces and all matter, that unifies QM and relativity, that resolves Einstein's Enigmas, and that reintroduces common sense into physics — this theory has largely remained a secret. Today, however, this is beginning to change as more physicists, like Nobel laureate Frank Wilczek, are telling people that nature is made of quantized fields.

The Core theory, which summarizes our best current understanding of fundamental processes, is formulated in terms of quantum fields. Particles appear as secondary consequences; they are localized disturbances in the primary entities – that is, in quantum fields. – F. Wilczek (W2008, p. 236)

<u>The plan</u>. The next five chapters will take you through the evolution of QFT step by step. The five force fields of nature are described in Chapters 2-5, beginning with gravity, the oldest known force, and ending with the two most recently understood ones: the weak field and the Higgs field. The two matter fields are then described in Chapter 6.

Chapters 7-9 describe Einstein’s enigmas and how they are resolved by QFT. In Chapter 7 you will see that the paradoxes of special relativity are natural and understandable consequences of the way fields behave. In Chapter 8 you will learn that gravity is not caused by curvature of space-time; it is a field like the other fields, and space and time are fundamentally different, in accord with our natural perceptions. In Chapter 9 you will see how QFT resolves the wave-particle duality problem, and provides a simple solution to the “Schrödinger’s cat” dilemma, also known as the *measurement problem*.

Chapter 10 (“The triumph of QFT”) provides an overview of QFT and its many successes, among which are a simple understanding of Einstein’s famous $e = mc^2$ equation (not found elsewhere, to my knowledge) and what Wilczek called “one of the greatest scientific achievements of all time” (that few people know about). It also shows that two of these fields are made out of more basic fields called quarks and gluons. Finally, out of fairness, the gaps in QFT are also described.

What’s missing. This is not a physics textbook. QFT is a complex subject, and this book only skims the surface, sometimes to the point of oversimplification. My aim is to convey the essence of what quantum fields are and how they behave without burdening you with the details and the mathematics. Although to physicists these details are important and the mathematics essential, they would cause anguish to many readers, or at least cause them to lose interest. (I once read that for every equation, you lose a thousand readers.) While the mathematics of QFT is part of its beauty, QFT can be understood, or at least appreciated, at a conceptual level without equations. As Paul Dirac once said,

> **Mathematics is only a tool and one should learn to hold the physical ideas in one’s mind without reference to the mathematical form. – P. Dirac (D1958, p. viii)**

There are also many things happening in physics today that are not included. You will not read here about black holes, supernovas, or the origin or fate of the universe. Cosmology is a fascinating topic, but it is not relevant to an understanding of the laws of nature. Nor will you read about attempts (e.g., superstring theory) to explain why fields have the properties they do. Why and wherefore are also fascinating questions, but QFT does not provide answers. Another noteworthy gap is consciousness, a mysterious process that happens behind our very noses, that neither QFT nor any other theory can even begin to explain.

Common sense. In this book I will frequently invoke common sense, but I must clarify what I mean by the term. I believe that common sense is choosing, wherever possible, the simplest, most intuitively-satisfying explanation *that is consistent with our observations*. This is not the same thing as accepting intuition blindly. An early example occurred when evidence that the earth is round triumphed over our intuitive belief that it is flat. A more recent example occurred when evidence for atoms dispelled our intuitive notion that matter is solid continuous stuff. If evidence, even very subtle evidence, requires that we change our long-standing intuitive view, we do it *because of*, not in spite of, common sense. However, the new view must make sense; it must be understandable, even if it wasn't what we expected. "OK, the earth is a round ball and not a flat pancake; I can live with that", we said. Or "OK, matter is made of tiny atoms held together by electric forces", or even, "OK, electrons and protons are not particles, they are little blobs of field. I can live with that." The job of science is to look beyond our intuition and to give up old ways of thinking when the evidence requires it. That's what I call common sense.

However, there is one intuitive concept that I will retain at all costs: that there is a reality. Those who don't believe that, or who are not interested in what it consists of, should not read this book. Those who do read the book will find that QFT offers a picture of reality that may not be what we expected, but one that we accept because of our common sense.

Features. References to the Bibliography are by author's initial and year of publication, e.g., E1923. There are also many references to speeches by Nobel laureates that can be found at www.nobelprize.org — the best single place I know to learn physics. Footnotes are used for comments that can safely be ignored by the casual reader. I have made abundant and possibly excessive use of quotes, both to add authority to my claims (I am not an expert in QFT) and also because I think it is important to read the words of the scientists who created the theories.

Now if you're thinking that physics = *hard, difficult, impossible*, please don't be discouraged. It is not easy to change our intuitive picture of matter from something that exists *in* space to an abstract property *of* space. Yet you will see that there is nothing paradoxical or inconsistent about QFT – it's just different from what you're used to. You may even find (like me) that the field picture of nature is more philosophically satisfying than the particle one. On top of this, QFT rests on mathematics that are the most beautiful equations I have ever seen or could imagine. If you will read ahead with, as Einstein once said, “a fair amount of patience and force of will”, you will achieve a basic understanding of the only consistent, non-paradoxical theory of what the world is made of: *quantum fields*.

Good luck.

CHAPTER 2

GRAVITY — The “blue” field

> **That one body may act upon another at a distance, through a vacuum, without the mediation of anything else… is to me so great an absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it. — Isaac Newton (B1966, p. 319)**

The story of the gravitational field is the oldest and longest story in physics. When Isaac Newton created the first theory of gravity in 1666, the concept of a field did not exist and, despite Newton’s misgivings (above), gravity was thought of as *action at a distance*. In other words, it was believed that one object (e.g., the earth) exerts a force upon another object (e.g., an apple or the moon) without anything in between. Newton, in his refusal to accept this concept, showed a remarkable insight and premonition of the field concept that was to come centuries later.

It wasn’t until the 20th century that gravity was given full status as a field. Today we believe that there is a gravitational field — a condition of space — that extends outward from every object in the universe and exerts a force on other objects. Gravity is a *force field*, as distinct from the *matter fields* that we will encounter in Chapter 6, but compared with other force fields, gravity is very weak. It takes the mass of the entire earth to hold a scrap of paper onto a tabletop, and yet that paper can be easily lifted by the small electric charge generated by rubbing a small comb. But before going further, I must clarify what is meant by mass.

MASS

Many people confuse mass with weight, among them my favorite cynic Ambrose Bierce, author of The Devil's Dictionary:

> **Gravitation. The tendency of all bodies to approach one another with a strength proportional to the quantity of matter they contain - the quantity of matter they contain being ascertained by the strength of their tendency to approach one another. This is a lovely and edifying illustration of how science, having made A the proof of B, makes B the proof of A. – A. Bierce (B1958, p. 51)**

Mass is inertia. Bierce's statement, although funny, is wrong. The quantity of matter they contain, i.e., their mass, is *not* ascertained by the strength of attraction between the bodies. In physics, mass means the tendency of an object to resist changes in course or speed. It is what physicists call *inertia*. Inertia is a kind of "I don't like to be pushed" feeling. It represents the difficulty in getting something going and the difficulty in stopping it once it is moving, or even the difficulty in deflecting it from its path. If you don't believe in inertia, try pushing a grand piano and a toy car and see which is easier. Even when we learn that "objects" are really made of fields (Chapter 6), the role of mass will remain essentially the same. Some fields are harder to change than other fields, and the difference is caused by a mass term in the field equations.

Mass is not weight. It is *weight* that corresponds to Bierce's "tendency of all bodies to approach one another". Weight is the pull of gravity — the downward force exerted by the earth. If there is no gravity there is no weight, but there still is inertia. In ordinary speech, the two concepts are often blurred, as in "a heavy object is harder to push", but in physics, weight is the pull of gravity and mass is inertia. They are different concepts.

Fig. 2-1. When John Godina, 1997 world champion shot putter, puts the shot, he must overcome its mass, or inertia, which is not the same thing as weight. (photo BBC Sport)

Consider, for example, a shot putter (Fig. 2-1). When John Godina picks up the 16 pound (7.2 kg) shot, he is working against gravity, which is not much of a problem for him. But then he must propel the ball forward, and that’s where the real effort comes in, because he is then working against the inertia of the ball. Because its inertia is much greater than a baseball’s, he is not able to make it go nearly as fast. Of course after he releases it gravity comes into play again and pulls it back to earth (after traveling, in the case of Godina, some 21.8 meters). In short, gravity and weight are involved during the lifting and trajectory phases, but it is mass or inertia that is the important factor in the throwing or putting phase.

To illustrate the difference, consider the same thing happening on the moon, where the force of gravity is 1/6 that on earth. When Godina picks up the shot he will find that it only weighs about 2½ pounds! “Oh,” he might think, “this shot put will be a piece of cake.” However, when he starts to propel the ball forward, he will be in for a surprise; he will find himself pushing against the same *mass* as on earth and he will only be able to impart to it the same velocity that he did on earth! Of course the weaker gravity again becomes a factor during the trajectory phase — and Godina will be pleased to see the shot travel farther, but the resistance he feels during the “putting” process, and therefore the imparted velocity, are the same as on earth, despite the lighter weight. If you can grasp this, you have understood the important difference between gravitational attraction (weight) and mass (inertia). Saying that gravity depends on something so basic and universal as mass is not a mere tautology; it is an experimental fact that makes gravity a very special and unique field — a fact that later became the basis for Einstein’s theory of gravity.

The story of how the gravitational field entered into physics is a strange one, involving apples and moons, predictions and confirmations, problem-plagued expeditions, and the two greatest scientific geniuses of all time. And it might not have happened if not for the Great Plague.

NEWTON'S GRAVITY

Science before Newton. In 1665 the physical sciences were very young, consisting basically of two areas of knowledge. One area lay in the heavens where, thanks mostly to Copernicus (1473-1543) and Johann Kepler (1571-1630), the sun was recognized as the center of the solar system with six known planets traversing elliptical orbits around it. (An ellipse is a kind of squashed circle.) The telescope had been invented and features could be distinguished on the moon. The other area of knowledge had to do with moving objects on earth, especially falling objects. Galileo (1564-1642) showed that as objects fall toward the earth, they accelerate at an equal rate, regardless of their mass (neglecting air resistance). There were other discoveries, of course, but it would not be far wrong to say that "elliptical orbits" in the heavens and "falling bodies" on earth constituted most of scientific knowledge at the time. And to accomplish even that was not easy. It required the overthrow of centuries-old Greek science, including Plato's insistence that heavenly bodies move in perfect circles about the earth, and the teaching of Aristotle (who was often wrong about physics) that heavier bodies fall faster than light bodies.

It is worth noting the different ways in which these understandings were achieved. Galileo, of course, was able to see with his own eyes the behavior of falling bodies, but Copernicus, who originated the heliocentric theory, could not see the sun's centricity directly, nor could Kepler see the elliptical orbits. All that could be seen is the apparent position of the sun and planets in the sky, and these could be explained in several different ways. Copernicus' theory was adopted because it provided the simplest explanation for the observations, without the complicated "epicycles" and "eccentrics" (we might call them curlicues) that were needed in the geocentric theory. But even Copernicus needed a few curlicues, and when Kepler showed that elliptical orbits fit the data without them, his theory was adopted because it offered the simplest explanation. These are examples of common sense: choosing the simplest explanation that can account for the observations, even if it means discarding an old belief. The further we go in understanding nature, the less we can see directly and the more we must rely on theoretical constructs that are removed from our perceptions. It is common sense that tells us to choose the explanations that are the simplest and require the fewest assumptions.

Good as Kepler's theory was, there still remained a nagging question: *Why* do the moon and planets travel in elliptical orbits? What is there about ellipses that made nature choose them instead of, say, circles or squares? Could someone perhaps find a simple law that could explain those ellipses?

Isaac Newton (1642-1727). In 1666, 50 years after the death of Shakespeare, the great fire hit London. One square mile was reduced to ashes and 13,000 houses were destroyed. The fire also ended the Great Plague that had devastated the city the previous year, killing almost 20% of its population. King Charles II himself had fled to the countryside to escape, as did a certain 22-year-old scholar who had just received his bachelor's degree from Cambridge University. This young man had a very unhappy childhood. His father, a farmer who couldn't sign his own name, died before he was born. His mother remarried and left him in the care of grandparents with whom he did not get along. Whether scarred by his childhood or through genetic flaws, the youth developed into a most strange and eccentric man,

> **…solitary, joyless, prickly to the point of paranoia, famously distracted… and capable of the most riveting strangeness… Once he inserted a bodkin into his eye socket and rubbed it around just to see what would happen… At least half his working life was given over to alchemy and wayward religious pursuits… [But] set atop these odd beliefs and quirky traits… was the mind of a supreme genius.. — B. Bryson (B2003, p. 46)**

The name of the young man was Isaac Newton.

Thoughts on falling objects. Newton had helped out on the farm before but in 1666, thanks to the plague, he had time for thinking. This is when he began his work on optics that brought him early fame, and this is when he invented the branch of mathematics known as *calculus*. But the achievement for which he is best-known had its origin one day as he watched an apple fall to the ground. Newton wondered if the same mechanism that caused the apple to fall might also hold the moon in its orbit about the earth. Could it be that the moon, like the apple, is pulled toward the earth? Could that explain why it follows an elliptical path around the earth? Could it be that planets are similarly pulled toward the sun and that this would explain their elliptical orbits?

To those who haven't studied physics, it may not be clear how the moon can be compared to a falling apple. The apple falls to earth, but the moon seems to be always at the same height (or distance above earth), without appearing to fall. What, then, is the similarity? Well, suppose we take an apple and throw it horizontally as hard as we can. It will not fall to earth immediately, but will continue in its trajectory, gradually getting closer to the ground until it hits it. Or if we fire a bullet horizontally it can travel for hundreds of meters before gravity finally pulls it to earth. If there were no gravity, both the apple and the bullet would continue forever in the direction in which they were pushed; it is gravity that alters the straight line path and brings the object back to earth.

Now let's talk about really fast speeds. Suppose that a rocket is launched vertically (to get it out of the atmosphere) and is then rotated 90°, after which a second stage is fired. One of three things can happen. (a) If the second blast is weak, the rocket will start to move in a horizontal direction but, like the apple and the bullet, will lose altitude and eventually fall to the ground. (b) At the other extreme, if the blast is very strong, the rocket will travel so fast that the earth's gravity won't have enough time to pull it down and it will move farther away, heading toward outer space. (c) A third possibility arises if the engine blast is strong enough to prevent the rocket from falling to earth, but not strong enough to enable it to escape into outer space. In that case, the rocket will keep going around the earth. The downward pull of gravity stops it from flying off into outer space, but it isn't strong enough to bring it down to earth. Thus the rocket keeps "falling" in a never-ending revolution around the earth. These possibilities are illustrated in Fig. 2-2.

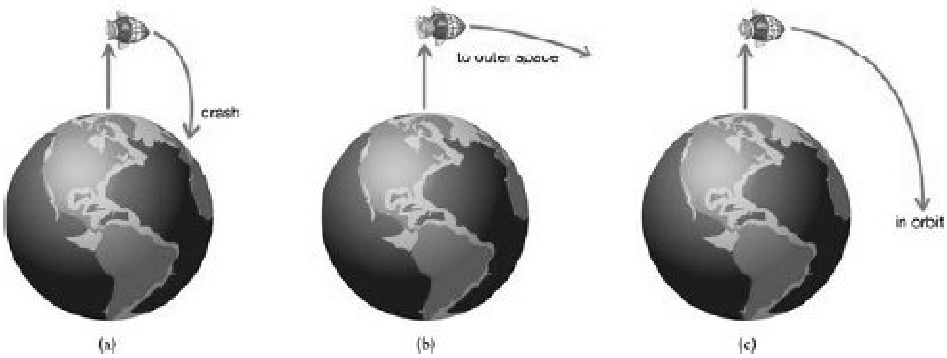


Fig. 2-2. Three rocket flights. (a) The blast is not strong enough to overcome the earth’s gravity and the rocket falls to earth. (b) The blast is so strong that the rocket escapes the earth’s gravity and travels into outer space. (c) The blast is of intermediate strength and the falling is just enough to keep it in a stable orbit.

A new mathematics. If one is a scientist, not just a philosopher, it’s not enough to have an idea; the idea must be tested by comparing it with reality. Could the earth’s gravity explain why the orbits of the moon and the planets are ellipses? Would it predict 29.5 days for the lunar month?[1] Unfortunately, there was no available mathematical method that could answer that question, so Newton did what he had to do: he invented calculus. He also had to know how the force of gravity varies with distance, and for this he made the natural assumption of an *inverse square law*. Natural, that is, if one assumes that the force is diluted in proportion to the extra area it has to cover (Fig. 2-3).

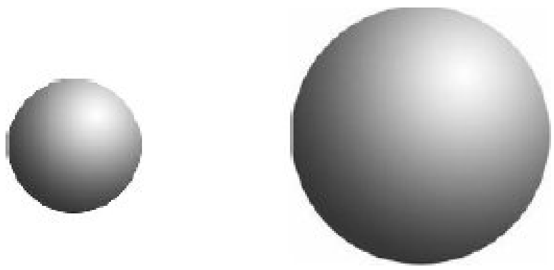


Fig. 2-3. The inverse square law. The radius of the sphere on the right is twice the radius on the left but its surface area is four times greater,[2] so the gravitational field is spread out over four times as much area.

Lo and behold, the result was… *elliptical trajectories!* But that was not enough. The next step was to calculate the moon’s period of revolution, and for that Newton needed to know the size of the earth and the distance to the moon, both of which had been measured in ancient Greek times.

Calculational difficulties. As it happened, Newton’s result did not match the actual lunar period. It is not known what caused the problem,[3] but it was enough to make him put the matter aside while he went on to make a name for himself in optics and other fields of physics. Among many accomplishments, he invented a new type of telescope using mirrors instead of lenses, which he presented to his one-time fellow refugee, King Charles II. His discovery about gravity might still be lying on a shelf if Sir Christopher Wren, the architect of St. Paul's Cathedral, had not offered a prize.

Eighteen years later, in 1684, Wren was dining with two friends and the discussion turned to astronomy. Wren offered 40 shillings for an explanation of Kepler's elliptical orbits. Edmond Halley, of "Halley's comet" fame, was one of the diners and he decided to consult with Newton:

> **Dr. Halley… asked him what he thought the curve would be that would be described by the planets supposing the force of attraction toward the sun to be reciprocal to the square of their distance from it… Sir Isaac replied immediately that it would be an [ellipse]. The Doctor, struck with joy and amazement, asked him how he knew it. "Why," saith he, "I have calculated it." – A. DeMoivre (B2003, p. 47)**

Newton then told him the whole story, including the disagreement with the lunar period. Halley urged him to try again. This time Newton refined his methods, and also had a better value of the earth's radius to work with. According to one report, when he saw that the answer was finally going to come out right, he grew so excited that he was forced to stop and let a friend continue for him. When Halley saw Newton's result he was equally excited. He not only induced Newton to write up his work, but paid for the publication. The result was the Principia Mathematica which, despite being in Latin and made "intentionally difficult by Newton so that he wouldn't be pestered by mathematical 'smatterers'" (B2003, p. 48), is generally considered the greatest scientific work ever written.

Acclaim. Newton soon became the most acclaimed scientist the world had known. Galileo had discovered regularities about falling bodies and Kepler had shown that the planets move in elliptical orbits around the sun, but it was Isaac Newton who made the great synthesis, explaining both of these observations on the basis of a single universal force.

> **Newton had matched the Greeks at their grandest and defeated them. [He produced] an overall scheme of the universe far more elegant and enlightening than any the ancients had devised. And**

> **the Newtonian scheme was based on a set of assumptions, so few and so simple, developed through so clear and so enticing a line of mathematics that conservatives could scarcely find the heart and courage to fight it. It excited awe and admiration among Europe's scholars… For the eighteenth century at least, man gloried in a new intellectual optimism that he had never experienced before and has never experienced since. — I. Asimov (A1982, p. 152)**

Controversy. Newton's success was not without controversy. When he tried to publish his work, Robert Hooke, Curator of Experiments for the Royal Society, claimed priority with respect to some of the ideas:

> **I have often wondered why the planets should move about the sun according to Copernicus's supposition, being not included in any solid orbs… nor tied to it, as their centre, by any visible strings… nor yet move in straight lines, as all bodies that have but one single impulse ought to do?… [A] cause of inflecting a direct motion into a curve may be from an attractive body placed in the centre; whereby it continually endeavours to attract or draw it to itself. For if such a principle be supposed, all the phenomena of the planets seem possible to be explained by the common principle of mechanic motions. – R. Hooke (I2002, p. 83)**

Hooke had even demonstrated his idea with a large pendulum hung from the ceiling. Although Hooke was "a nasty, argumentative individual, antisocial, miserly, and quarrelsome [who] maintained a lifelong enmity [toward Newton], clearly founded on jealousy" (A1982, p. 144-149), there was something to what he said, so Newton inserted a reference to Hooke, saying that he was now going to expound on Hooke's ideas in greater detail.

Mathematics and physics. Newton marked a turning point in physics. Before Newton mathematics and physics were largely unrelated disciplines. One did not have to be a gifted mathematician to work out the laws of nature. After Newton this was no longer the case, with a few exceptions like Michael Faraday (see Chapter 3). In fact, Newton's calculus was just the beginning. Later generations saw partial differential equations, Riemannian geometry and Hilbert algebra become an inextricable part of physical theory.[4] It goes without saying that this increased reliance on mathematics has contributed toward the gap between science and public understanding.

Where's the field? Astute readers will have noted that, for all of Newton's success, there was one thing missing. I'm referring to the subject of this chapter, the gravitational field. At the time, the discovery of the law of gravity was enough. It explained virtually all that was known in a way that everyone could understand, even if they couldn't do the math. That was enough to satisfy the most curious intellect and no one thought to ask how the force was transmitted. No one, that is, except Newton. This remarkable man, who had uncovered the secret of gravity, could not believe in *action at a distance* (see quote at the head of the chapter). He clearly wanted there to be a gravitational field, but finding no evidence he avoided speculation. And so scientists continued to believe that one body acts upon another at a distance for another 200 years, until Faraday and Maxwell developed the concept of the electromagnetic field. And it would be 50 more years before Einstein developed the field equations that gave gravity its "full citizenship" in the family of fields.

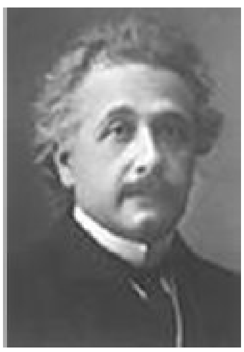
EINSTEIN'S GRAVITY

When Einstein began to tackle gravity in 1907, science had come a long way from the days of Newton. Not only had electricity and magnetism been discovered, but the concept of an electromagnetic field had replaced action-at-a-distance and was on the verge of replacing the ether. This led naturally to gravity also being seen as a field. After all, why would nature use a field for one force and action at a distance for another?[5] There had also been great advances in astronomy, partly due to Newton's new telescope. Two more planets had been discovered, with only Pluto remaining to be found (and later "demoted"). However, there still were no gravitational field equations – the *sine qua non* of a true field. There also was a small problem pertaining to the orbit of the planet Mercury.

A problem with orbits. The science of astronomy had become so precise that a French astronomer, Urbain Le Verrier, was able to uncover a slight irregularity in the orbit of Mercury, the planet closest to the sun. He found that the orbit changes in such a way that its point of closest approach to the sun (the *perihelion*) advances by a tiny amount each year. When I say tiny, I mean only 43 seconds (about 1/100 of a degree) per *century!* However, because it could not be explained, this 1/100 of a degree loomed as large as a barn door. As Le Verrier stated:

> **[There must be] some as yet unknown action on which no light has been thrown… [This is] a grave difficulty, worthy of attention by astronomers – U. J. J. Le Verrier (P1982, p. 254).**

And so Einstein, like Newton, was confronted with an unexplained observation about orbits and, like Kepler's elliptical orbits, this small irregularity would lead to a major revolution in scientific thought.



Albert Einstein (1879-1955). Albert Einstein didn't like school. He was once told by a high school teacher "You will never amount to anything, Einstein". After dropping out of high school, he applied to the renowned Swiss Federal Polytechnic School (ETH) in Zurich, but failed the entrance examination. Knowing that admission was granted to anyone with a Swiss high school diploma who passed the "Matura" examination, Einstein enrolled at a cantonal school in Aarau, where he passed the Matura with flying colors. He did fairly well at ETH with the help of a friend, Marcel Grossman, who let him copy his lecture notes. In 1901 he obtained his diploma. Failing to find a suitable academic position, partly because of the anti-Semitism of the time,[6] he obtained a job (again with the help of friend Grossman) at the patent office in Bern. In 1905, his *annus mirabilis* (miracle year), he published four papers that contained major contributions in three different areas of physics, including the quantum paper that was to win him the Nobel Prize (see Chapter 3). Two years later, while still working at the patent office, he began to consider the problem of gravity. His approach was quite different from that of Newton, but like Newton's, it started with a thought about falling objects.

Thoughts on falling objects. Einstein's thought was not about apples. His falling object was a person, and he saw him fall only in his mind's eye:

> **When working on a comprehensive paper on the special theory of relativity… there occurred to me the happiest thought of my life…. *for an observer falling freely from the roof of a house there exists –* at least in his immediate surroundings – *no gravitational field* (Einstein's emphasis). Indeed, if the observer drops some bodies then these remain relative to him in a state of rest or of uniform motion, independent of their particular nature… This simple thought made a deep impression on me. It impelled me toward a theory of gravitation. – A. Einstein (P1982, p. 178-179)**

Of course this was only a thought-experiment, but today most of us have seen images on television of astronauts in orbiting spacecraft that illustrate what Einstein saw in his mind's eye — what we call *zero gravity* or *zero-g* (Fig. 2-4). If there had been spacecraft in Einstein's day, everyone might have had the same happy thought!

Fig. 2-4. Astronaut Jim Bagian enjoying zero gravity in a spaceship as a result of two opposing effects: gravity that pulls him (and the ship) toward the earth, and inertia, which manifests itself as an apparent upward (centrifugal) force. The two effects cancel.

Whether it is a falling man or an orbiting satellite, the effect of inertia is to create an apparent outward force — the same force we feel when riding in a car that goes around a tight curve. This inertial effect is equal and opposite to gravity and therefore cancels the pull of gravity. In physics language, the *gravitational mass* and *inertial mass* are equal. This is not a tautology, as Ambrose Bierce thought, but recognition that the pull of gravity is proportional to the inertia of the object being pulled. Einstein called this the "Principle of Equivalence", and it became the basis for his new theory of gravity that he called the *general theory of relativity*.

A new mathematics. Like Newton, Einstein realized that he needed a new kind of mathematics to develop his theory. The problem was that the gravitational field is not uniform; it varies from point to point. Einstein needed a mathematical technique that could handle this point-by-point variation, but unlike Newton, he was unable to invent it. Instead he did the next best thing: he consulted his faithful mathematician friend, Marcel Grossmann. "Grossmann," he wrote, "you must help me or else I'll go crazy!" (P1982, p. 212). Grossmann helped. He found that such a method, called Riemannian geometry, was already in the literature. This enabled Einstein to make a good start on the problem, but it was only a start:

> **The Einstein-Grossman paper published in 1913 contains profound physical insight into the nature of measurement, some correct general relativistic equations, some faulty reasoning, and clumsy**

notation. — A. Pais (P1982, p. 216) [7]

<u>**Calculational difficulties**</u>. Newton's challenge was to obtain agreement with the 29½ day lunar month. Einstein's was to obtain agreement with the 43 seconds of angle for Mercury. Like Newton, he failed at first, but unlike Newton he didn't put the matter aside. He kept at it obsessively, never losing confidence that his Principle of Equivalence was the key. He kept at it at the University of Zürich, where he received his first academic appointment. He kept at it during a year in Prague, and he kept at it at after taking a position at the University of Berlin in 1914. Max Planck, perhaps the most eminent physicist before Einstein, said to him:

> **As an older friend I must advise you against it. In the first place, you will not succeed; and even if you succeed, no one will believe you. — M. Planck (P1982, p. 239)**

Einstein did not take Planck's advice. During his long struggle he took more wrong turns and dead ends than a drunk driver on Saturday night, and he published every one. Einstein joked about this in a letter to Paul Ehrenfest:

> **That fellow Einstein suits his convenience (E*s ist bequem mit dem Einstein*). Every year he retracts what he wrote the year before. – A. Einstein (P1982, p. 250)**

The truth is, Einstein's grasp of Riemannian geometry was a bit shaky. David Hilbert, a leading German mathematician who heard him lecture in 1915, said (with a touch of exaggeration):

> **Every boy in the streets of Göttingen understands more about four-dimensional geometry than Einstein. — D. Hilbert (S1986, p. 188)**

<u>**Success at last!**</u> On Nov. 25, 1915, Einstein announced that he had finally succeeded in explaining "the secular rotation of the orbit of Mercury discovered by Le Verrier… without the need of any special hypothesis". This was at least as great an accomplishment as Newton's, and it almost gave Einstein a heart attack:

> **This discovery was, I believe, by far the strongest emotional experience in Einstein's scientific life, perhaps in all his life. Nature had spoken to him. He had to be right. 'For a few days, I was beside myself with joyous excitement.' Later, he told Fokker that his discovery had given him palpitations of the heart… when he**

saw that his calculations agreed with the unexplained astronomical observations, he had the feeling that something actually snapped in him. – A. Pais (P1982, p. 253)

The years of searching in the dark for a truth that one feels but cannot express, the intense desire and the alternations of confidence and misgiving until one breaks through to clarity and understanding are known only to him who has himself experienced them. – A. Einstein (P1982, p. 257)

Einstein's theory provided a mathematical basis for the gravitational field and vindicated Newton's intuitive rejection of action-at-a-distance. In addition, the theory was consistent with his Principle of Relativity (see Chapter 7), which says that the laws of nature are identical in all moving systems, regardless of their rate of motion provided it is constant. The gravitational field could now take its place alongside the electromagnetic field as a property of space, with its own field equations.

Controversy. Just as Newton had a last minute dispute with his nemesis Hooke, so Einstein had a last minute dispute with David Hilbert. In June of 1915, before Einstein made his final breakthrough, Hilbert invited him to give two lectures at Göttingen. On November 20, five days before Einstein presented his final theory, Hilbert submitted a paper that contained virtually the same equations, albeit without the vital calculation of Mercury's orbit. Einstein complained angrily to Hilbert that he had plagiarized the ideas that had been presented at Göttingen.[8] It was only after Hilbert apologized, saying "this talk had completely slipped my mind," that Einstein could write:

I have struggled with complete success against a feeling of bitterness… I think of you once again with untroubled friendliness and ask you to try to do the same regarding me. – A. Einstein (P1982, p. 260).

Predictions. The agreement with Mercury's orbit was enough in itself to convince people that Einstein's new theory of gravity was valid, but Einstein didn't stop there. He went on to make several predictions, chief of which was that light rays passing close to a massive object would be deflected from their original path by the gravitational pull of the object. Einstein even pointed out that this prediction could be tested by looking for shifts in the position of stars during a solar eclipse. (Without the eclipse, stars whose position is sufficiently close to the sun would not be visible.) As it happens, Einstein was not the only one to make such a prediction. The other person was (pause)

Isaac Newton:

> **I shall conclude with proposing only some Queries, in order to a farther search to be made by others. Query 1. Do not bodies act upon light at a distance, and by their action bend its rays; and is not this action strongest at the least distance? – I. Newton (N1979, p. 339)**

Newton's speculation was based on his belief that light is corpuscular, but the actual deflection could not be calculated at the time because the speed of light was not known. When it was finally measured,[9] the calculation was made and it was found that Newton's theory predicted a deflection of 0.9 angular seconds. Coincidentally, this was the same value found by Einstein in the early forms of his theory, but it had increased to 1.8 angular seconds in the final formulation. Thus astronomers not only had the task of determining if there is a deflection, but which theory it agreed with.

Expeditions. Sir Arthur Eddington pointed out that a suitable eclipse was slated to occur on November 6, 1919, so the Royal Astronomical Society of London dispatched teams of astronomers to Brazil and West Africa to make the necessary observations. Both expeditions ran into bad weather and technical problems; the African group produced only two usable images while the Brazil group obtained 7 usable images out of 26 exposures. When the groups returned and analyzed the data, they found that there was indeed a deflection. Since no one had observed *any* deflection before, this was in itself a triumph. More importantly, they found that the magnitude agreed with Einstein's prediction. As the physicist J.J. Thomson said,

> **The deflection of light by matter, suggested by Newton in the first of his Queries, would itself be a result of first-rate scientific importance; it is of still greater importance when its magnitude supports the law of gravity put forward by Einstein. – J.J. Thomson (P1982, p. 305)**

Coincidentally, in 1735 France had sent two teams to prove (or disprove) a result of Newton's theory: the earth's equatorial bulge.[10] One expedition went to Peru and the other to Scandinavia. They encountered even more mishaps than the 1919 expeditions, but in the end, Newton's theory was confirmed, French scientists were probably disappointed, and the equatorial bulge is now an accepted fact.

Acclaim. A successful prediction is worth a dozen retroactive explanations, and the drama of the expeditions only added to the excitement. The result was that Einstein was virtually canonized.

> **Einstein was now world famous. Ordinary people might not understand his theories and might only grasp dimly what it was all about but there was no question that they understood him to be *the* scientist. No scientist was so revered in his own time since Newton.**
> **– I. Asimov (A1982, p. 677)**

However, this recognition did not carry over when Einstein was awarded the Nobel Prize in 1921. His theories of relativity were specifically excluded from the citation as being "the subject of lively debate in philosophical circles". Instead the award was given for his 1905 paper on the quantum nature of light. Einstein showed his feeling about this evasion by delivering his acceptance lecture on relativity theory, without mentioning the work for which he had been cited.[11]

Relativistic twins? When one looks at the paths that Newton and Einstein followed while pursuing their theories of gravity, one is struck by the many similarities: the unexplained data on orbits, the sudden insight about falling objects, the need for a new mathematics, the calculational difficulties, the retroactive agreements, the controversy, the problem-plagued expeditions, and the final triumph and acclaim. Both men had worked in the same eccentric and lonely way, divorced from other scientists, armed with a great feeling of self-reliance while struggling with new concepts and difficult mathematics, and both produced earth-shaking results. One can't help but wonder if these two greatest of scientists, born 237 years apart, were "relativistically related", conceived as twins in some ethereal plane in a far-off galaxy and sent to earth to solve a matter of some gravity.

THE GRAVITATIONAL FIELD

The concept of a field does not come easily. For centuries people thought of space as mere emptiness with objects embedded in it. Stars, mountains, molecules, electrons — these are all things that exist in space; if nothing is present, there is only empty space. In other words, space is a *container*. In the field picture, on the other hand, there is no empty space. Fields pervade space; they are a condition or property *of space*. You can't have space without fields. Fields obey laws that specify how a change at one point affects the field at adjacent points and how the change is propagated through space.

> **The action of the earth on the stone takes place indirectly. The earth produces in its surroundings a gravitational field, which acts on the stone and produces its motion of fall… The intensity and direction of the field at points farther removed… are thence determined by the law which governs the properties in space of the gravitational fields themselves. — A. Einstein (E1961, p. 72)**

Propagation speed. An important feature of fields is that changes in intensity do not propagate through space instantaneously[12] but proceed, point by point, with a speed determined by a number c in the field equations. This number is about 300,000 km/sec (186,000 miles/sec), and it is the same for all fields. Since light is a field, you will not be surprised to learn that c is the speed of light. If an object were to suddenly appear in space, the gravitational field generated would first appear near the object and then extend outward with velocity c until all of space acquires, to some extent, the attractive property created by that object.

Gravity-induced contraction. One of the consequences of Einstein's equation is that gravity causes objects to contract. This is not surprising when you consider that gravity interacts with and affects the EM fields that hold things together (cf. Lorentz contraction in Chapter 7). In fact, gravity interacts with all other fields. The contraction is very small, but it was critical in the detection of…

Gravity waves. If gravity is a field then, just as an oscillating electron in an antenna sends out radio waves, so a large mass moving back and forth will send out gravitational waves. Einstein predicted gravity waves a few months after he found the equations for his general relativity theory and, although there was no direct evidence for such waves at the time, the analogy with EM waves was so compelling that most physicists accepted their existence. What many people don't know is that Einstein later did a flip-flop.

Einstein flip-flops. In 1936, Einstein and his assistant Nathan Rosen found a mathematical result that was inconsistent with gravity waves. This led them to submit a paper to the *Physical Review* entitled "Do gravity waves exist", with a conclusion saying that the answer is "no", or perhaps just questioning their existence. (The paper no longer exists.) The editor sent the paper for review by Howard Percy Robertson, a noted expert in the new science of cosmology, who found an error in the mathematics. The article was then returned to Einstein with Robertson's (anonymous) critique. Einstein, who had a disputatious streak (see footnote 8), was furious and sent the following letter to the editor:

> **Dear Sir: We had sent our manuscript for publication and had not authorized you to show it to specialists before it is printed. I see no reason to address the — in any case erroneous — comments of your anonymous experts. On the basis of this incident, I prefer to publish the paper elsewhere. Respectfully, Albert Einstein ("Einstein vs. the Physical Review" in *Physics Today*, Sept. 2005, p. 43)**

The "elsewhere" that Einstein chose was the *Journal of the Franklin Institute* in Philadelphia, but before it could be published, Robertson met with Einstein's new assistant, Leopold Infeld, and showed him the error. Infeld agreed, and went to Einstein to point out the error (which Einstein claimed he had also detected the night before). Einstein quickly revised the paper, keeping the title "Do gravity waves exist?" but changing the conclusion from "no" to "yes". This paper was published in the Franklin Institute journal in 1937. Einstein never submitted another paper to Physical Review.

Evidence. The first evidence for gravity waves came from a pulsar 30,000 light years from the earth. A *pulsar* ("pulsing star") is an extremely dense star that emits periodic microwave (EM) radiation. The pulsar in question was discovered on July 2, 1974, and given the strange name PSR1913+16 (pronounced "Joe"). The important thing about this particular pulsar is that it is part of a doublet — that is, it has a companion star — and the two revolve around each other with a period of 8 hours. Now this revolution is the kind of motion that, as we saw above, creates gravity waves, and these waves carry away energy. As the stars lose energy, they will move closer together and revolve at a faster rate, just as a spacecraft in a low orbit has a shorter period than one in a higher orbit.

Now the rate of revolution for a pulsar doublet can be measured by observing the shift in the microwave radiation as the emitting star alternately approaches and recedes. (This is essentially the same shift that we hear in a train's whistle as it approaches and then recedes from a crossing.) Joseph Taylor and Russell Hulse, who discovered the pulsar, monitored its period for 20 years. They not only found a decrease, but the amount agreed with the predictions of Einstein's field equations. For this feat, Taylor and Hulse were awarded the 1993 Nobel Prize in physics.

In 2015 gravity waves were detected directly at LIGO (Laser Interferometer Gravitational-Wave Observatory) by measuring the contraction of one arm of an apparatus, as compared with another arm at right angles. The arms were 4 km long and the laser beam was made to traverse the paths almost 500 times, giving a total travel distance of over 1000 km. This experiment captured the imagination of the public and will stand as one of the great feats of experimental physics, alongside the famous Michelson-Morley experiment of 1887, which it resembles (see fig. 7-1). In fact, when you read the description of the M-M experiment you will see that the LIGO experiment is not as hard to understand as you thought. The field nature of gravity and the existence of gravitational waves were now established beyond any doubt.

Fig. 2-5. The earth is shown surrounded by a "blue" gravitational field. The field that exists within the earth is not shown.

Color it blue. An analogy used throughout this book is the use of color to visualize fields that in themselves are too abstract to picture. Thus I ask you to picture in your mind's eye a blueness around the earth that represents its gravitational field (Fig. 2-5). The blue is most intense at the surface and becomes fainter at greater distances. The choice of blue, of course, is arbitrary. Later we will add more fields and colors but, unlike an artist's colors, these colors remain separate and do not combine, the way blue and yellow combine to give green. Each field retains its own identity and nature, regardless of what other fields are present.

Where's the Q? You may have noticed that something is missing. If this book is about Quantum Field Theory, you ask, then where is the quantum? The answer is found in Chapter 8, where you will learn that, contrary to what you may have heard, QFT *is* consistent with general relativity. In fact there is a natural place for it, and the gravitational quantum has even been given a name: *graviton*.

> **The graviton is unknown, as yet, to experimental science. Nevertheless, we shall accept it and its conjectured properties as the proper starting point for the theory of gravitational phenomena, just as the photon with its attributes initiates the theory of electromagnetic phenomena. The evidence for the existence of the graviton is indirect, but impressive. – J. Schwinger (S1990, p. 81)**

So for now we can forget about the Q, and we should be grateful. The Q will appear in the next chapter and will turn out to be an important aspect of the electromagnetic field.

SUMMARY

Newton’s insight that the moon is a falling object led to a theory of gravity that explained the elliptical orbits of the moon and planets. Although Newton did not treat gravity as a field, he strongly rejected the concept of action-at-a-distance. The idea that gravity is a field was introduced in 1915 by Albert Einstein. After many false starts, Einstein succeeded in deriving a gravitational field equation from his Principle of Equivalence, which states that gravitational forces are indistinguishable from inertial forces. This theory is known as the General Theory of Relativity. The success of general relativity in explaining the discrepancy in the orbit of Mercury and in predicting the deflection of starlight (confirmed a year later) caused Einstein to achieve a world-wide fame that surpassed even that of Newton. One consequence of Einstein’s theory is the existence of gravity waves, which were detected for the first time in 2015.

CHAPTER 3

ELECTROMAGNETIC FORCES - The "green" field

> **[My purpose is] to explain the action between distant bodies without assuming the existence of forces capable of acting at sensible distances. The theory I propose may therefore be called a theory of the Electromagnetic Field. — J. C. Maxwell (P1982, p. 229)**

The field concept was first introduced into physics in 1845 by Michael Faraday as an explanation for electric and magnetic forces. Later, in 1864, the concept was given a mathematical basis by James Clerk Maxwell, who combined Faraday's two fields into a single *electromagnetic* (EM) field. (It would be another 50 years before Einstein developed a field theory for gravity.)

Like gravity, the EM field is a force field, but unlike gravity it does not interact with everything. It is created by objects that carry an electric charge and it exerts a force on objects that carry an electric charge. In other words, electric charge is to the EM field what mass is to the gravitational field. However, while everything has mass, electric charge is found only in some objects (or fields). Further, unlike mass, electric charge comes in two varieties that are called, quite arbitrarily, *positive* and *negative*. The particles that make up ordinary matter are the positive proton, the negative electron, and the neutron, which carries no electrical charge. (We will continue to call these entities particles until Chapter 6, where their field nature will be revealed.) In the normal atom the number of electrons equals the number of protons, so there is no net electrical charge. However, some electrons are loosely bound to the nucleus so a net charge can be created by stripping away electrons, as happens when you run a rubber comb through your hair.

As its name implies, the EM field is a complex field that contains both electric and magnetic components. At first, electric and magnetic forces were thought to be different and unconnected, but it was eventually learned that they are really two aspects of the way electric charges interact. A charged particle at rest creates an electric field and a charged particle in its vicinity will feel a force because of the electric field. A charged particle in motion (or an electric current) creates a magnetic field and a nearby moving charge will feel a force because of this magnetic field. It is for this reason that the two fields can be combined into a single electromagnetic field.

Magnetism. For a long time the origin of magnetism was unknown. This led Ambrose Bierce to offer the following definitions in his “Devil’s Dictionary”, written in the late 19th century:

> **Magnet. Something acted upon by magnetism.**
> **Magnetism. Something acting upon a magnet.**
> **The two definitions immediately foregoing are condensed from the works of one thousand eminent scientists, who have illuminated the subject with a great white light, to the inexpressible advancement of human knowledge. – A. Bierce (B1958, p. 84)[1]**

At the time, Bierce’s cynicism was justified. However, after the discovery of the electron in 1897 it became apparent that magnetism is caused by the motion of electrons. The connection between moving charges and magnetic force is seen, for example, in an ordinary electric motor, where electric currents exert a magnetic force that makes the rotor move. Even in the case of the earth’s magnetism, it is circulating electric currents in the earth’s core that create its magnetic field, which in turn makes compasses possible. But this is not a physics textbook, and we don’t have to go into any of this. All you need know is that electric and magnetic fields are two aspects of the way charged objects (or fields) exert forces on each other.

Force fields are directional. The EM field is complex in another way. Force fields are directional; when you push something, you push in a direction. Gravity is an attractive force and its direction is toward the object that created the field, but the EM force is not necessarily attractive. The electric field created by a positive charge exerts an attractive force on a negative particle but a repulsive force on a positive charge (“like charges repel, unlike charges attract”). The magnetic field is even more complicated; it exerts a force on a moving charge in a direction perpendicular to both the direction of motion and the direction of the field. The important thing to remember is that the components that make up the EM field have directions associated with them.

Color it green. Following our use of color as a means to visualize fields, I will choose green for the EM field. Thus, in your imagination you should picture space as having two colors, blue and green, corresponding to the gravitational and electromagnetic forces that are present. The “blue” gravitational field is strongest near massive bodies and the “green” EM field is strongest near charged bodies, but both fields are found everywhere, even if their intensity is very weak. The fields created by an atomic nucleus and by the earth are depicted in Fig. 3-1. However, when using the color analogy we must keep in mind that it only shows the existence of a field; it does not indicate its complexity or direction.

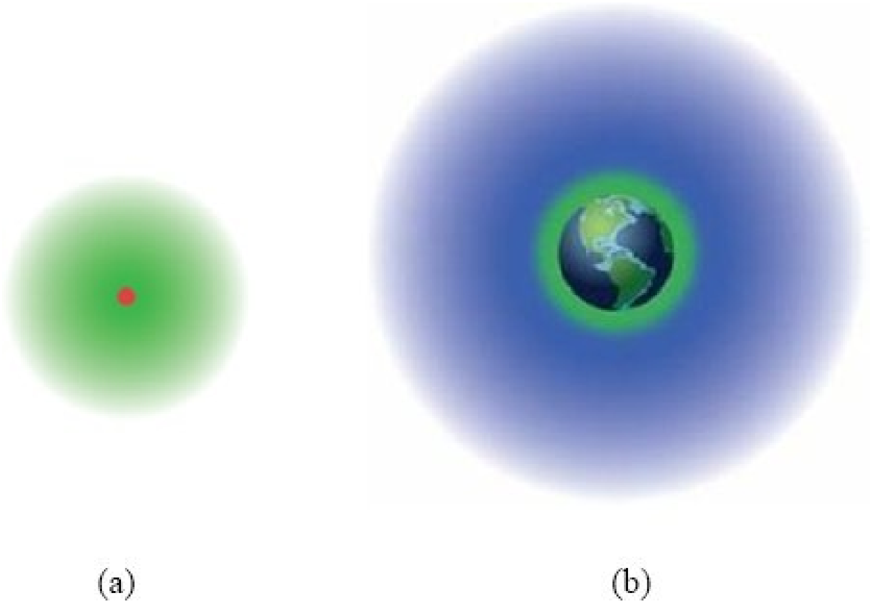
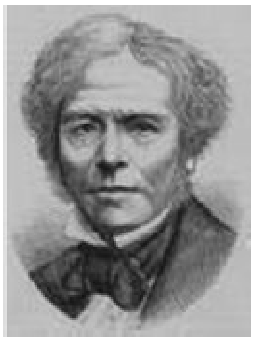


Fig. 3-1. (a) The “green” EM field around a positively-charged atomic nucleus (gravitational field not shown). (b) The earth with its “blue” gravitational field (cf. Fig. 2-5) and “green” EM field created by the earth’s magnetism.

The story of how the EM field entered into physics is a two-part story, with the centennial year 1900 forming an appropriate dividing line between the two parts. The theory developed in the 19th century is called *classical*, while the 20th century saw the development of the quantum theory. We will begin, of course, with…

THE CLASSICAL EM FIELD

Because of their complexity, you will not be surprised to learn that the various laws of electricity and magnetism were discovered gradually over a period of 100 years or more. Among the many scientists making contributions were Benjamin Franklin (1706-1790), Henry Cavendish (1731-1810), Charles Coulomb (1736-1806), André Ampère (1775-1836), Hans Oersted (1777-1851), Carl Friedrich Gauss (1777-1855), and Joseph Henry (1797-1878). We will pick up the story a bit later, when a most unusual scientist became the first person to use the word *field*.



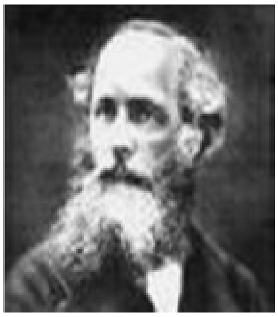
Michael Faraday (1791-1867). Michael Faraday was the son of a blacksmith, one of ten children, who started out as an apprentice to a bookbinder. He was and remained virtually illiterate in mathematics, but this didn't stop him from pursuing his keen interest in science and becoming one of England's (and the world's) foremost scientists. He was at least partly responsible for the invention of the electric motor, the transformer, and the generator – three staples of our modern electrified society. However, his greatest contribution from our standpoint is that he introduced the field concept into physics.

Lines of Force. Faraday came to the idea of fields after observing the way iron filings line up when placed on a sheet of paper over a magnet (see Fig. 1-1). This is an experiment that every physics student performs today "as a matter of course." Take a bar magnet, place a card over it and sprinkle iron filings on the card. Then jiggle the card and watch how the filings arrange themselves into curved lines in the space between and around the magnetic poles.

Faraday believed that these lines, which he dubbed "lines of force", must indicate a physical reality. *Something* must be there, he thought. On Nov. 4, 1845, he created physics history by writing the word *field* in his notebook for the first time.[2] Faraday later extended the field concept to interactions between electric charges, although he could not make the electric field manifest itself as vividly. To Faraday, these fields were almost tangible agents that carry force from one charged object to another.

> **As a result of the more careful study of electromagnetic phenomena, we have come to regard action at a distance as a process impossible without the intervention of some intermediary medium. If, for instance, a magnet attracts a piece of iron, we cannot be content to regard this as meaning that the magnet acts directly on the iron, through the intermediate empty space, but we are constrained to imagine – after the manner of Faraday – that the magnet always calls into being something physically real in the space around it, that something being what we call a 'magnetic field'. In its turn this magnetic field operates on the pieces of iron, so that the latter strives to move towards the magnet. — A. Einstein (E1966, p. 71)**

But Faraday's idea, like Newton's, was only conceptual – a guess, if you will. It was James Maxwell who gave it a firmer basis.



James Clerk Maxwell (1831-1879). Maxwell was appointed professor of physics at Aberdeen in 1856, the same year that Faraday turned 65. Inspired by Faraday's ideas and possessing the mathematical ability that Faraday lacked, Maxwell went on to develop the first field theory in physics. He took the laws that had been discovered about electricity and magnetism up to that time and converted them into field equations. He even introducing a new law of his own.[3] The result was a unification of all electric and magnetic laws into a set of four field equations that were presented to the Royal Society in 1864, and are now known as Maxwell's equations.

> **A new concept appears in physics, the most important invention since Newton's time: the field. It needed great scientific imagination to realize that it is not the charges nor the particles**

> **but the field in the space between the charges and the particles that is essential for the description of physical phenomena. The field concept proved successful when it led to the formulation of Maxwell's equations describing the structure of the electromagnetic field. – A. Einstein and L. Infeld (I2007, p. 92)**

A new mathematics. To do this, Maxwell, like Newton and Einstein, needed a new type of mathematics. Fields are properties of space and therefore are described by equations that are fundamentally different from the ones Newton used to describe the motion of objects. For an object, it is enough to give its location and velocity plus an equation stating how the location and velocity change with time. Fields, on the other hand, require a specification of intensity *at every point of space*, plus an equation describing how that intensity changes. Further, this change must depend only on what is happening at that point. (This is what physicists mean by the word *local*, as opposed to action-at-a-distance.) Again quoting Einstein:

> **There is a profound difference between the theoretical representation of ponderable bodies and Maxwell's theory of EM processes. Whereas we may consider the state of a body as being determined by the positions and velocities of (to be sure) a very large but finite number of atoms and electrons, we must now use continuous spatial functions to specify the electromagnetic state, so that a finite number of parameters cannot be considered as sufficient." – A. Einstein (B1966, p. 544)**

Fortunately for Maxwell, he did not have to invent the mathematics himself, as Newton did. The method, known as *partial differential equations*, had been invented by the French mathematician Augustin Cauchy while Faraday was evolving his intuitive concept of fields. Maxwell, who was highly skilled at mathematics, used the new technique to reformulate the laws of electromagnetism into a simple and elegant set of equations that describe how the field intensity at one point influences what happens to the field at adjacent points.[4]

Waves. These equations led Maxwell to predict the existence of EM waves that propagate through space, as a change in field intensity at one point affects the intensity at adjacent points. For example, when electrons move back and forth in an antenna they create a corresponding alternating fluctuation or oscillation in the surrounding EM field that propagates outward, with changes at one point engendering changes at neighboring points. It is like water waves traveling through water or sound waves traveling through air, except that the speed is determined by properties of the EM field, not by properties of water or air. EM waves are also more complicated than these simpler waves because the propagation involves an interplay between the electric and magnetic components, with changes in one inducing changes in the other.

> **Suppose that at a point out in otherwise empty space a magnetic field changes… The changing magnetic field creates a changing electric field, which in turn regenerates a magnetic field. That is, as time elapses, at that point there is an oscillation between the two kinds of fields. In addition, the fields vary from one point of space to another. All this is reminiscent of a more familiar phenomenon, the motion of waves. For example, if you are floating in a boat on the ocean surface, the passage of ocean waves by that point in space is experienced in the course of time as a rhythmic up and down motion of the boat – it oscillates. And, at a given instant, if you look out, there in the distance – removed in space – are the successive troughs and crests of the advancing waves. — J. Schwinger (S1986, p. 13)**

Light. Maxwell had an even greater surprise awaiting him. From these equations, derived from the existing laws of electricity and magnetism, there emerged a *constant* that he called *c*, from the Latin word *celeritas*, meaning "swiftness". (A constant is a fixed number, as opposed to something that can take on different values, like field strength.) It is this number that determines the speed of propagation of EM waves, and it can be calculated from the constants in the electric and magnetic laws that were incorporated into his equations. We can imagine Maxwell's surprise and delight when *c* turned out to be equal to the known speed of light!

> **This velocity is so nearly that of light that it seems we have strong reason to conclude that light itself (including radiant heat, and other radiations if any) is an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws. – J. C. Maxwell (B1966, p. 343)**

And so in one fell swoop Maxwell unified the laws of electricity and magnetism, created a mathematical basis for fields, predicted EM radiation, and explained light, which had been an unsolved mystery for hundreds of years. One cannot overestimate the importance of this achievement.

> **The most fascinating subject at the time that I was a student was**

> **Maxwell's theory. What made this theory appear revolutionary was the transition from forces at a distance to fields as fundamental variables. The incorporation of optics into the theory of electromagnetism, with its relation of the speed of light to the electric and magnetic units… it was like a revelation… I cannot suppress the remark that the pair Faraday-Maxwell has a most remarkable inner similarity with the pair Galileo-Newton – the former of each pair grasping the relations intuitively, and the second one formulating those relations exactly and applying them quantitatively. – A. Einstein (B1966, p. 561)**

Other waves. Maxwell's phrase "other radiations if any" was perhaps the greatest understatement in the history of physics. In 1888 Heinrich Hertz of Germany, spurred by Maxwell's discovery, generated radio waves for the first time and the unit of frequency, abbreviated Hz, was named in his honor.[5] Radio waves range from frequencies as low as 3000 Hertz (1 Hz = 1 oscillations per second) to 1 gigaHertz (1 GHz = 1 billion oscillations per second). The higher frequencies are used for FM radio, television, cell phones, etc. In the mid 1900's this frequency range was extended to 300 GHz by the development of *microwave* technology — called "microwaves" because of their very short wavelength.[6] This 100-million-to-one range of frequencies, however, is just the beginning. After microwaves comes the radiant heat that Maxwell referred to, now called *infrared* radiation because its frequency range (300-300,000 GHz) lies just below that of red light. Then comes visible light, which occupies a tiny niche in the vast spectrum of EM radiation.

Digression on color. Before continuing with the EM spectrum, I would like to say a word about color and where it comes from. I'm not referring to the imaginary colors that we are using to help visualize fields; I'm referring to the colors that we actually see in the world around us. Color is determined by the frequency of radiation, with red being the lowest visible frequency, followed by orange, yellow, green, blue and violet, in that order. When all colors are present, as in sunlight, our eyes perceive the result as white. The component colors of white light can be demonstrated by passing it through a prism, or by looking at the rainbow created by nature's prism (water droplets).

However, and it's a big however, the colors we see do not exist outside of our minds. EM radiation in itself has no color; it is merely an oscillation in the EM field intensity. It is only when this radiation reaches our eyes that it starts a process that turns the oscillations into sensations of color. Our eyes contain three types of photosensitive molecule that respond to three different frequencies of light. (This is why display devices such as a TV screen are able to reproduce the full range of color using only three phosphors.) Each photosensor, in turn, is connected by a different neural pathway to the mysterious part of the brain where the sensation of color is created, but we are not aware of the separate intensities of the three neural signals. Our brain combines them to create a host of color sensations, depending on the amount of each component that is present. It’s a bit like a chef mixing different flavors to create a new taste, or a composer putting notes together to create a chord, except that a good chef can detect the individual elements in a mixture and a good musician can distinguish the component notes in a chord. This is not the case with color. The component colors do not survive; they are lost to the new color sensation produced by the combination. For example, if red and blue combine to produce purple, the red and blue hues can no longer be separately identified, although they are still present. Our brains process color very differently from the way they process other sensory signals.

If you don’t believe that colors are created in our brain, consider Fig. 3-2. Note the difference in perceived color between the central tiles on the front and top faces, even though the frequencies of radiation coming from them are the same.

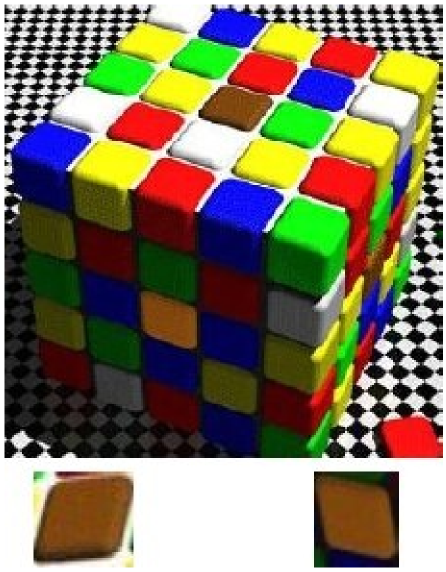


Fig. 3-2. The central tiles on the front and top faces are actually the same color, as can be seen when they are shown separately (below). The light coming from them has the same wavelength, but because of the difference in surroundings, our brain creates different colors for the two tiles. (R.B. Lotto and D. Purves, Nature Neuroscience 2:1010-1014, 1999)

<u>More waves</u>. Returning to the EM spectrum, after visible light comes *ultraviolet* radiation, so named because it is “beyond the violet”. This is what we seek when we go to a tanning parlor or become painfully aware of if we stay in the sun too long. At still higher frequencies are x-rays, discovered in 1895 by Wilhelm Roentgen (who won the first Nobel prize in physics in 1901), although their EM nature wasn’t demonstrated until 1912. Finally come gamma rays, the highest known EM frequencies, discovered by Henri Becquerel in 1896 as emanations from uranium, and later by Marie and Pierre Curie in 1899 as emanations from radium, for which discovery all three shared the 1903 Nobel Prize. The EM nature of gamma rays was demonstrated shortly afterward by Ernest Rutherford, who called them *gamma rays* to distinguish them from two other types of nuclear radiation that he called *alpha* and *beta*.

<u>The ether</u>. When the wave nature of light first became evident, a substance called ether was postulated “to do the waving”. This luminiferous (light-bearing) substance was assumed to pervade space everywhere, yet to be completely invisible and undetectable. Its sole function was to be the medium for the propagation of light. Despite many attempts to develop a mathematical theory for this mysterious substance, its physical properties remained elusive, but the concept didn't die. Even with the success of Maxwell's equations and the discovery of different types of EM radiation, the idea that a field could exist by itself was too hard for many physicists to accept. Just as water waves are oscillations in water, they thought, so there must be a substance to carry the oscillating electric and magnetic fields.

> **One knows where our belief in the ether stems from. When light is on its way to us from a far star… it is no longer on the star and not yet on the earth. It is necessary that it is somewhere, sustained, so to say, by some material support. – H. Poincaré (P1982, p. 127)**

Maxwell, however, never liked the ether, calling it "a most conjectural scientific hypothesis". To him it was enough to have waves propagated through the electromagnetic field, a la Faraday.

Goodbye ether. Einstein's theory of special relativity (see Chapter 7), introduced in 1905, was based on the idea that there is no such thing as absolute motion. Yet an ether, if it existed, would constitute a natural "at rest" system and thus would be a basis for defining absolute motion. Thus as the theory of relativity gradually became accepted, the concept of an ether receded. In QFT there is no ether; there is only space that has physical properties. Indeed, the whole question may be one of semantics. As Paul Drude wrote in 1900:

> **The conception of an ether absolutely at rest is the most simple and most natural – at least if the ether is conceived to be not a substance but merely space endowed with certain physical properties. – P. Drude (P1982, p. 121)**

Interference. Interference occurs when two fields converge at a point in space and either reinforce or cancel each other, depending on the direction of their forces. An example is the way wakes from two boats overlap and produce a pattern of peaks and troughs, depending on how they come together at a given point. Another example is shown in Fig. 3-3(a), which is a photograph of water waves passing through two apertures into a tank. The light and dark areas correspond to peaks and troughs in the water.

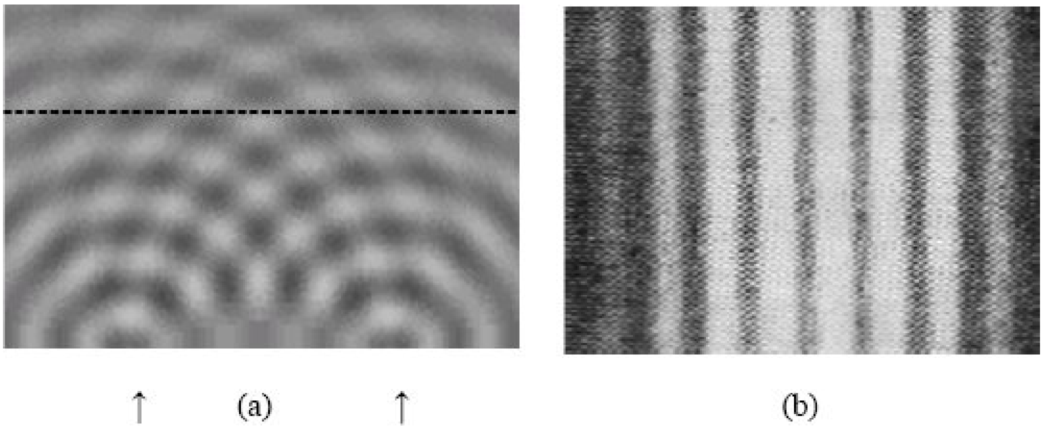


Fig. 3-3. (a) Photo of water waves emanating from two apertures in the wall of a tank (arrows). As the waves spread out they either reinforce or cancel each other, producing peaks and valleys (light and dark areas). If we look at the pattern along the dashed line we see that the light and dark areas alternate. (b) The same alternating pattern is seen when light reaches a photographic film after passing through two small slits, as the waves coming from each slit either reinforce or cancel each other like the water waves in (a). (H2007, p. 300)

Well, the same thing can happen with light. Fig. 3-3(b) shows an interference pattern created by light that has passed through two slits and then falls on a photographic film. The bright bands are regions where the electric (and magnetic) fields of the two light beams are pushing in the same direction and reinforce each other. The dark bands are dark because the two fields are pushing in opposite directions and cancel each other. The fact that light exhibits interference effects was demonstrated as early as 1803, and was the chief reason that Newton's corpuscular theory of light was abandoned.

A time for celebration. As the 19th century drew to a close there was a feeling of optimism and celebration in physics. The two known forces of nature, gravity and electromagnetism, were understood and the mystery of light was explained. Some leading scientists went so far as to claim that the age of discovery was over. All that remained, they felt, was to clean up a few "loose ends." However, one of those loose ends turned out to be a time bomb.

THE QUANTUM EM FIELD

Changing colors. The time bomb exploded on December 14, 1900, and it led to a revolution in physics unlike any that had occurred before. It all started because objects change color when they get hot. Consider, for example, a burning coal, or the element of an electric stove, or the filament in a light bulb. As the object gets hotter it begins to glow – first a deep red, then orange, then yellow, and if it gets hot enough it becomes white (hence the expression "white heat"). What is happening is that thermal energy in the object is being converted into EM radiation, and as the temperature increases, the frequency of the radiation increases. (Reminder: different frequencies of light create different color sensations in our brains.) The radiation extends over a range of frequencies, but its intensity reaches a peak at a frequency that depends on the temperature of the object. Fig. 3-4 shows graphs of radiation intensity vs. frequency for a particular kind of hot object called a *black body*.

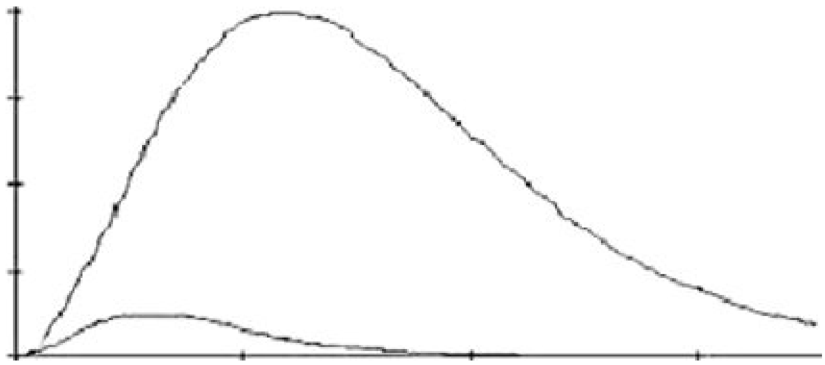


Fig. 3-4. EM radiation intensity (vertical scale) plotted vs. frequency (horizontal scale) for a "black body" at two temperatures: 5,000 °K (lower curve) and 10,000 °K (upper curve). (www.galileo.phys.virginia.edu)

The problem was that according to Maxwell's EM theory the energy radiated should not have a peak; it should be independent of frequency. Attempts to find *ad hoc* mathematical expressions that would describe the strange experimental result were only partially successful: One expression worked well at low frequencies, another at high frequencies, but no one could find an equation to fit the whole curve, much less a theory that would explain it. That is, until Max Planck came along.

<u>Max Planck (1858-1947)</u>. In 1900, Planck was one of Europe's most respected physicists, a professor at the University of Berlin — the same school that 13 years later would attract Albert Einstein to its hallowed ranks. Planck made two breakthroughs in regard to the radiation problem. First, he found a way to combine the two empirical equations into a single equation. More importantly, he found an explanation for why the equation worked.

<u>Enter the quantum</u>. Planck realized that his equation (called Planck's law) would make sense if the energy of the EM radiation, instead of varying continuously, is emitted in discrete amounts or "chunks" of energy. He called these chunks *quanta* (singular *quantum*) from the Latin word meaning "so much". Planck's concept can be understood by considering two bowls of sugar. One bowl contains granulated sugar, so that the coffee drinker is free to take as much or as little as she wants (as long as she doesn't try to subdivide a grain). The other bowl contains lumps of sugar, so the user can take only discrete amounts — one lump, two lumps, or (God forbid) three lumps. So it is with EM radiation: we may find one quantum of energy, two quanta, or more, but we cannot find any amounts in between. However, these quanta are not confined to a localized region like lumps of sugar; each is a spread-out unit of field with its own discrete energy.

<u>Success</u>. Planck immediately recognized the importance of his discovery. His son (who was later executed by the Nazis for involvement in the plot to kill Hitler[7]) remembered his father saying:

> **Today I have made a discovery as important as that of Newton — (B1966, p. 473)**

In his 1920 Nobel lecture Planck put it this way:

> **Here was something entirely new, never before heard of, which seemed called upon to basically revise all our physical thinking.**

How right he was! Planck had introduced discreteness into physics, and physics would never be the same.

Energy. The term *energy* may require a little explanation. Even though it occurs in ordinary speech as well as in physics, the meanings are not quite the same. To a physicist, energy is the ability to do work, and work means *pushing* or *pulling* — that is, moving something by exerting a force on it. Given this definition, it should not be hard to see that the EM field contains energy, since it is made of electric and magnetic fields that exert forces. The stronger the field, the more pushing it can do and hence the more energy it contains. Recognizing that fields contain energy is an important step in helping us see them as having real existence. (You may recall that it was the energy carried away by the gravitational field that was the first experimental evidence for gravitational waves.)

Frequency. The key assumption in Planck's theory is that the energy of each field quantum depends on its rate of oscillation, i.e., frequency. The constant that relates energy to frequency is called Planck's constant, written as the letter *h*. In case you're wondering how big *h* is, its magnitude is 6.6×10^{-27} (0.0000000000000000000000000066) ergs per Hz (reminder: 1 Hz = 1 oscillation per second). Now the erg (the unit of energy) itself is pretty small — there are 40 billion in a food Calorie — so you can see we're dealing with very small amounts of energy here. This relation between energy and frequency will be a key step in the derivation of $e = mc^2$ shown in Chapter 10.

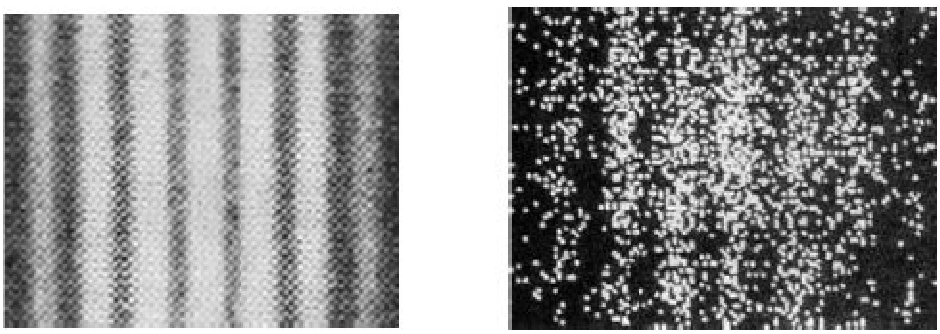
More predictions. Not only did Planck's theory explain the experimental data, it also enabled the calculation of three fundamental constants, including the charge of the electron.[8] As it happened, the value that Planck obtained for the electron charge disagreed with the best existing measurement by 35%, which probably caused Planck some anxious moments. However, there was so much right in his theory that the agreement was considered close enough. It would be another nine years before Robert Millikan obtained an improved value of the electronic charge that was in full agreement with Planck's calculation, confirming the theory beyond any doubt.

Enter Einstein. In 1905, his *annus mirabilis*, Albert Einstein carried the quantum revolution a step further – a "quantum leap", we might say (inside joke). Picking up where Planck left off, Einstein looked at the absorption of EM radiation in what is called the *photoelectric effect*, This effect, first observed by Heinrich Hertz in 1887, refers to the way electrons are ejected from a metal when ultraviolet light strikes it. It was clear that the energy of the absorbed light is transferred to the electrons, but increasing the intensity of light did not increase the energy of the ejected electrons — it only caused more electrons to be ejected. On the other hand, raising the frequency of radiation increased the energy but not the number of electrons. Einstein pointed out that this behavior is to be expected if the radiation is absorbed in the same discrete chunks, or quanta, that Planck postulated to explain the radiation spectrum. It was this achievement for which Einstein received the 1921 Nobel Prize, his theory of relativity being too controversial at the time.

Einstein's photon. Einstein's view of the EM quantum was not the same as Planck's. Einstein felt that if a quantum of light is emitted by a single atom at one location and then absorbed by a single atom at another location, surely it must be confined to a small region of space during its journey. That is, it must be a particle, as the great Newton had believed some two centuries earlier.

> **It appears to me that the observations on phenomena related to the generation and transformation of light can be understood better on the assumption that… the energy in a beam of light emanating from a point source is not distributed continuously over larger and larger volumes of space but consists of a finite number of energy quanta, localized at points of space, which move without subdividing and which are absorbed and emitted only as units. – A. Einstein (B1966, p. 544)**

Thus was born the concept of the *photon*,[9] symbolized by the Greek letter gamma (γ). However, Einstein still had to contend with the interference effects, which were shown to exist even at the single photon level (Fig. 3-5).



(a) (b)

Fig. 3-5. (a) The same two-slit interference pattern shown in Fig. 3-3(b). (b) The pattern when the light intensity is lowered. Each dot represents the detection of a single photon. The same pattern appears, showing that each photon interferes with itself, as only a spread-out field can do. (H2007, p. 301)

Fields vs. particles (Round 1). Thus was born the continuing battle between fields and particles. As Einstein said:

> **We are forced to conclude that… a single photon is responsible for the ability of the two beams to interfere, as well as for the absorption of light from one of the beams. It is evident that Maxwell's theory cannot account for this complex of properties of the photon. It does not provide us with any means to understand the atomistic character of the absorbed energy of radiation… [Yet] the interpretation of the photon as a pointlike structure does not admit of an explanation for the interference phenomena which are produced if both parts of the beam interact. – A. Einstein (1950 speech to International Congress of Surgeons, quoted in Physics Today, June 2005, p. 47)**

This is what two other great physicists had to say about the quandary:

> **There is in particular one problem… What becomes of the energy of a photon after complete emission? Does it spread out in all directions… Or does it fly out like a projectile in one direction…? In the first case, the quantum would no longer be in the position to concentrate energy upon a single point in space in such a way as to release an electron from its atomic bond, and in the second case, the main triumph of the Maxwell theory… would have to be sacrificed, both being very unhappy consequences for today's theoreticians." – M. Planck (Nobel lecture, 1920)**

> **Einstein was led to the formulation of the so-called 'hypothesis of light quanta', according to which the radiant energy, in contradiction to Maxwell's electromagnetic theory of light, would not be propagated as electromagnetic waves, but rather as concrete light atoms [i.e., localized photons], each with an energy equal to that of a quantum of radiation… However, the hypothesis of [localized] light quanta, which is quite irreconcilable with so-called interference phenomena, is not able to throw light on the nature of radiation. – N. Bohr (Nobel lecture, 1922)**

The dilemma led many physicists to the concept of wave-particle duality (see Chapter 9), or even to give up hope of comprehending reality. Einstein, however, could not do this. He believed there is a reality and that it must make sense. He spent the last half of his life searching for a unified field theory that would explain (among other things) the peculiar behavior of light quanta. Shortly before his death he wrote:

All these fifty years of pondering have not brought me any closer to answering the question, what are light quanta? – A. Einstein (P1982, p. 382)

Quantum collapse. The answer, according to QFT, is that Planck was right and Einstein was wrong. In QFT the photon *is* a spread-out field; the particle-like behavior occurs because each photon, or quantum of field, is absorbed as a unit. The answer to Planck's question, "how can the photon be in a position to "concentrate its energy upon a single point in space?" is simply that it isn't. It is spread out, but when it transfers its energy to an atom the entire field vanishes, no matter how spread-out it is. There is a big "whoosh" and the quantum is gone, like an elephant disappearing from a magician's stage. Quantum collapse will be discussed further in Chap. 10.

A minority view. I should warn you that most physicists today do not accept this view. They either follow Feynman in believing that photons are particles, or they believe in wave-particle duality, paradoxical as those views may be. And some, like Stephen Hawking, just don't worry about the problem (see p. 4). To me, the persistence of and insistence on a belief in particles or in wave-particle duality, despite the phenomenal success of QFT (see Chapter 10), is itself a bit of a paradox.

Self-fields. It is important to understand the difference between EM fields that surround charged objects as in Fig. 3-1(a) and photons that fly off into space. The former are *self-fields*; they are created by a source and can never leave that source. Field quanta, on the other hand, have a life of their own; they can go anywhere they like. So in addition to the "green" EM fields that surround charged objects and currents, we must also picture blobs of green that are radiated from various sources and spread out in space as they travel.[10] If one of these blobs collapses, the greenness suddenly disappears as the energy is transferred to an absorbing atom.

Another new mathematics. When Planck's quantum principle of discreteness was incorporated into EM theory, once again a new mathematics was needed. Newton had to invent calculus to solve his problem, while Maxwell and Einstein found existing mathematical techniques to do the job. The developers of QFT were also lucky, as a mathematical method for describing quantities that can take only discrete values already existed. It is known as *Hilbert algebra* or *Hilbert space*, after the German mathematician David Hilbert. In Hilbert space, the physical states of a field are described by complex *vectors*, and there are *operators* that transform one vector into another. It is these operators that appear in the field equations, not the field intensities themselves. However, a description of Hilbert algebra is even farther beyond our scope than partial differential equations, so you'll have to take this on faith. You can understand the results without understanding the mathematics that leads to them.

Spin. Finally, I must mention an important attribute of quantum fields called *spin* or *helicity*. Spin is the more common term, but it gives the false impression of an object spinning. In QFT spin or helicity is a measure of the internal complexity of the field, i.e., the number of internal components. The possible spin values of quantum fields are 0, 1/2, 1 or 2 (in units of Planck's constant divided by 2π). The EM field is fairly complex with two component fields (electric and magnetic), each of which is directional, and its spin is 1. You will be surprised to learn that the gravitational field is the most complex of all the fields, with a spin of 2. After all, you ask, what could be simpler than a single attractive force, but when Einstein introduced his field equations, the internal complexity expanded greatly. I didn't mention this in Chapter 2 because I thought you had enough to worry about, and anyway the reason for it is beyond the scope of this book. I ask you to accept that the gravitational field is more complex than the EM field and has a spin or helicity of 2 without asking why, or even understanding what that really means. Another thing I must ask you to accept on faith is that this abstract definition of spin, related to field complexity, causes a quantum to exhibit angular momentum just as a spinning particle would.

SUMMARY

The electromagnetic field, like gravity, is a force field; it is the agent that transmits force from one electric charge to another. It consists of two component fields, electric and magnetic. Unlike gravity, which is always attractive, the EM force can be attractive, repulsive, or even sideways. Its field nature was suggested by Michael Faraday in 1845 and field equations were developed by James Maxwell in 1864. These equations predicted EM waves (oscillations in field intensity) that travel at a speed of 300,000 km/second, which turned out to be the same as the speed of light. Not only did Maxwell's equations thus explain the entire field of optics, they also predicted a vast spectrum of EM radiation that includes radio and TV waves, microwaves, infrared, visible light, ultraviolet, x-rays, and gamma rays.

The quantum nature of the EM field was discovered by Max Planck while studying radiation from a hot object. He found that the data made sense if the radiation is not infinitely divisible, but consists of discrete amounts of energy that he called quanta. These quanta are now called photons. Planck also found that the energy of a photon is proportional to its frequency of oscillation.

Later Einstein showed that absorption of radiation occurs in the same discrete amounts of energy, suggesting that the quanta are actually particles. This was the first round in the fields-vs.-particles battle which goes on even today. In QFT a photon is a field and only a field. It spreads out in space and exhibits interference effects. When a photon is absorbed, no matter how spread out it may be, its energy is deposited into the absorbing atom. This is called *quantum collapse*.

The internal complexity of a quantum field is described by a number called *spin* or *helicity*. The gravitational field is the most complex of all fields, with a spin of 2 Planck units. The EM field, with its component electric and magnetic fields, has spin 1. The two fields introduced so far are summarized in the following table:

Field	**Spin**	**Quantum**	**Interacts with**	**"Color"**
Gravity	2	Graviton	Everything	blue
EM	1	Photon (γ)	charged fields	green

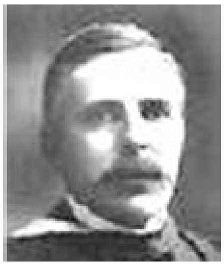
CHAPTER 4

THE STRONG FORCE - The "purple" field

> **Interactions between the [nuclear] particles can be described by means of a field of force, just as the interaction between the charged particles is described by the electromagnetic field. — H. Yukawa, 1935 (B1966, p. 1419)**

The early 1900's were an exciting time for physics. The "outer world" was well understood, but the "inner world" – the world of the atom - was just beginning to emerge. Because of this, physics had to make a major change in its methods. Before 1900, physics dealt with macroscopic things – things one could see and measure, like falling bodies, planets, moons and magnets. As the new century dawned, it became evident that there is a microscopic world filled with things so tiny that we can never hope to see or experience them directly. Physicists had to learn new ways to deal with these inaccessible entities — to take the slenderest and most indirect of clues and infer from them the nature of atoms and subatomic particles. It was by this process that a Japanese physicist was able to discover a new field of force: the so-called *strong force* that holds the atomic nucleus together. The story began in 1911 with the discovery of the nucleus by a man who came from a farm in New Zealand.

THE ATOMIC NUCLEUS



Ernest Rutherford (1871-1937). After obtaining his M.A. and B.Sc. degrees from the University of New Zealand, Ernest Rutherford was digging potatoes on his father's farm one day when he received news of a scholarship award from Cambridge University. He is reported to have said "That's the last potato I'll dig" and left immediately for England. Rutherford, like Faraday before him, was not highly skilled in mathematics, but had a great physical intuition and a gift for experiment. When he arrived in England, he was faced with one of the great mysteries of physics: what does the atom look like? (One theory depicted the atom as a solid round ball with electrons embedded in it like raisins in a fruitcake.)

Rutherford's atom. Rutherford's great feat was to determine the structure of atoms experimentally. To do this he used a piece of radium that emits a form of radiation that he called *alpha rays* (see "Radioactivity" in Chapter 5). By aiming this radiation at a thin sheet of gold foil, and analyzing the scattered radiation with the aid of a little mathematics (something Faraday would not have been able to do), Rutherford was able to determine that gold atoms are not solid balls, but are in fact mostly empty space with a very small positively-charged nucleus at the center.

> **One day in 1911 Rutherford, very sprightly and happy, came into [Hans] Geiger's laboratory and informed Geiger that he knew what the atom looked like… Geiger began his crucial experiments to test Rutherford's analysis on that same day and thus verified one of the greatest contributions to physics. This must be considered as the very peak of Rutherford's research and is certainly to be counted among the greatest of all scientific achievements. — H. Boorse and L. Motz (B1966, p. 705)**

The picture of the atom that emerged from this work is shown in Fig. 4-1. It consists of negatively-charged electrons circling a positively-charged nucleus, held in their orbits by electrical attraction, just as the planets are held in orbit by the sun's gravity.

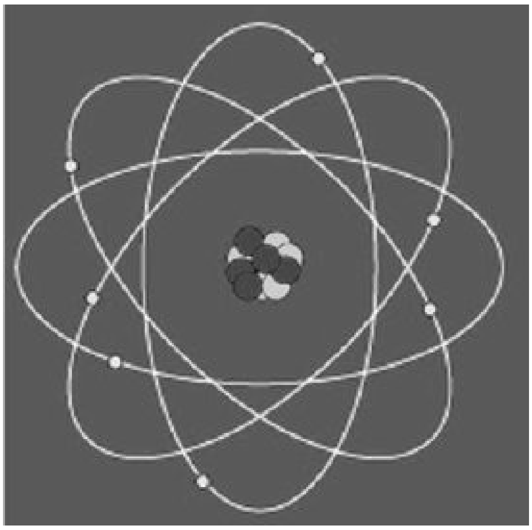


Fig. 4-1. Rutherford's atom consists of a nucleus made of protons and neutrons surrounded by orbiting electrons. If the sketch were to scale, the nucleus would be so tiny (100,000 times smaller than the electron orbits) as to be invisible.

The nucleus, it was soon discovered, is made of two types of particles: *protons* that carry positive charge and *neutrons* that carry no charge. These particles are collectively called *nucleons*. (Warning: This picture, which is often seen in books and articles, is not the picture provided by QFT. We will see in Chapter 6 that electrons and nucleons are not particles, but quanta of fields.)

History of atoms. Atoms had been known, or at least hypothesized, since the time of ancient Greece. The idea was first suggested by Democritus (470-380? BC), "the most successful of the Greek natural philosophers in the uncanny accuracy of his ideas" (A1982, p. 12). It was Democritus who coined the name *atom*, meaning "uncuttable". A Greek engineer named Hero even found experimental evidence to support the atomic theory.[1] Unfortunately, the atomic theory was rejected by Aristotle — in my opinion the most unsuccessful of the Greek natural philosophers in the uncanny *inaccuracy* of his ideas, who believed that everything on earth is made of four elements: earth, air, fire and water. Since Aristotle's writings became the scientific Bible of the middle ages, atomic theory was ignored by most scientists until 1803, when the chemist John Dalton revived it to explain the known chemical properties of the elements.

How big are atoms?

> **Atoms are tiny – very tiny indeed. Half a million of them lined up shoulder to shoulder could hide behind a human hair… A typical paramecium is about two microns wide, 0.002 millimeters, which is really very small, [but] atoms exist on a scale of minuteness of another order altogether. To get down to the scale of atoms, you would need to take each one of those microns and shave it into ten thousand finer widths… It is a degree of slenderness way beyond the capacity of our imaginations, but you can get some idea if you bear in mind that one atom is to a millimeter as the thickness of a sheet of paper is to the height of the Empire State Building. – B. Bryson (B2003, p. 134)**

The inimitable Mr. Bryson also said that the number of atoms in a gram of hydrogen "is equivalent to the number of popcorn kernels needed to cover the United States to a depth of nine miles, or cupfuls of water in the Pacific Ocean… It is a big number."

Well, if the atom is this tiny, what can we say about the nucleus, which is 100,000 times smaller? Following Bryson's analogy, I might say that if the atom were magnified to the size of the USA, the nucleus would correspond to a large building somewhere in Kansas.

The proton. The proton was discovered in 1918 by Rutherford, who gave it its name, meaning "first particle". The number of protons in a nucleus is called the *atomic number*. Since atoms are electrically neutral, the atom must contain the same number of electrons. For example, the hydrogen atom (atomic number 1) has 1 proton and 1 electron; the helium atom (atomic number 2) has 2 protons and 2 electrons, etc. The largest naturally-occurring atom is uranium (atomic number 92), with 92 protons and 92 electrons. The "job" of the proton, which is about 2000 times heavier than the electron, is to use its electrical attraction to keep the electrons from flying away.

The question now to be addressed is what stops the electrical repulsion between protons from causing the protons to fly away from each other? Clearly there must be another force in the nucleus that is strong enough to overcome the electrical repulsion. But first, we must bring into the picture…

The neutron. The neutron is the other fundamental particle in the nucleus. Its mass is slightly greater than the proton's, but it carries no electric charge. We might call it an "unattractive" sibling of the proton (moan). As a nucleon it contributes equally to the strong binding force, but its role in the nucleus goes beyond that. It also acts as a kind of spacer to separate the protons and thereby weaken the electrical repulsion between them. Without this weakening, the strong force would not be able to hold the nucleus together.[2]

Take the helium nucleus, which contains two protons and two neutrons. Without the two neutrons, the strong binding force would not be able to overcome the electrical repulsion between the protons and the nucleus would fly apart. With the neutrons, however, the binding force is increased because there are now four nucleons, while the electrical repulsion is weakened because the protons are not as close together. The result is a stable helium nucleus, also known as an alpha particle — the same radiation used by Rutherford in his gold foil experiment.[3] In nuclei with more protons, even more neutrons are needed to dilute the increasingly large repulsive force. This is why in most atoms the neutrons outnumber the protons. However, if there are too many neutrons the nucleus will break down because of the *Pauli Exclusion Principle* (see Chap. 6).

A delicate balance. In summary, nuclei are held together by a delicate balance among (a) the attractive strong force between nucleons, (b) the electrical repulsion between protons, (c) the dilution of this repulsive force by neutrons, and (d) the Exclusion Principle that prevents neutrons from "piling up". For a given number of protons there is a narrow range of allowable neutron numbers: not enough and the protons will fly apart because of electrical repulsion; too many and the nucleus will be unstable because of the Exclusion Principle. Atoms with the same atomic number but different numbers of neutrons are called *isotopes*; they have the same chemical properties and differ only in mass. Isotopes are labeled by the total number of nucleons they contain. For example, uranium with 92 protons can have between 142 and 146 neutrons in its nucleus, corresponding to isotopes called uranium 234 and uranium 238.

Finding the neutron. Because it is electrically neutral, the hunt for the neutron was not an easy one. It wasn't until 1932 when Rutherford's former student, James Chadwick, obtained proof of its existence. In an interesting footnote to history, it has been speculated that if the neutron had been discovered earlier, Germany might have succeeded in developing an atomic bomb before the US (B1993, p. 143), in which case World War II would have ended quite differently.[4] When one considers that Chadwick, the discoverer of the neutron, spent four years during World War I interned in Germany as an enemy alien, a certain irony emerges.

THE STRONG FORCE FIELD

Now let us return to the force that holds the nucleons together in the nucleus. The only forces known in 1934 were gravity and electromagnetism, but gravity is not strong enough and the EM force produces repulsion, not attraction. Clearly there must be a new force acting within the nucleus, and it must be strong enough to overcome the electrical repulsion. It is called, therefore, the *strong nuclear force*, or just the *strong force.*

A short-range force. The real puzzle about the strong force was not its existence, but its short range. If it fell off according to the inverse square law, like gravity and EM forces, it would be observable outside the nucleus — but it is not. A further indication of its short range is the fact that the binding energy (amount of energy needed to break up the nucleus) decreases in heavy nuclei. This shows that the binding forces between nucleons are not cumulative, and this can only happen if the range is small compared with the size of the nucleus.

Because of the success of Maxwell's EM field theory and the acceptance of gravity as a field, there was no doubt in anyone's mind that the strong force is transmitted by a field, but no one was able to come up with the right field equations to describe this new short-range force. The answer came from an unexpected quarter.

Enter Japan. Up to then, theoretical physics had been mostly a European activity, dominated particularly by Germany and Great Britain. The United States was still lurking on the edge and other areas of the world were completely dormant. As for Japan, it had maintained its ancient culture in isolation from the West until 1853 when Commodore Perry and his US Navy warships entered Yokohama harbor. This invasion led to the fall of the existing shogunate and the Meiji restoration. During the ensuing "golden age", many Japanese went to Europe to learn about Western science and Westerners were brought to Japan to teach. Never has a country accomplished so much so fast, and today Japan is one of the leading countries in the world in both science and technology.

Hideki Yukawa (1907-1981). Hideki Yukawa, son of a geology professor, graduated from Kyoto University in 1929. Five years later, as a resurgent militarism was engulfing Japan that would culminate in the invasion of China in 1937 and the attack on Pearl Harbor in 1941, Yukawa came up with the answer to the nuclear force quandary that had stumped Western scientists for over 20 years. The Yukawa field, as it is sometimes called, joined the ranks of the gravitational and EM fields as the third fundamental force field in nature.

The Yukawa field. Starting from the analogy with Maxwell's equations, Yukawa developed a field equation that contains a new term that changes the force from a slowly-varying one, like EM and gravity, to one that falls off rapidly — in fact, exponentially. The new term contains a constant that he called λ (lambda), and by choosing λ properly, Yukawa was able to make the theoretical range of the strong force match the experimentally-determined value. Yukawa thereby did what Maxwell had done for EM forces. He started with a force that was understood less well than EM forces and devised field equations to describe that force.

Color it purple. I choose purple as the color to represent the strong field. Around every proton and every neutron I ask you to picture a small purple halo (Fig. 4-2), representing the strong field. Unlike the green of the EM field and the blue of gravity, this halo does not extend very far, but nevertheless it is there. In Chapter 6 we will learn that the orbiting electrons shown in Fig. 4-1 are really fields, and this will add "yellow" to the space around the nucleus. These fields all exist independently and do not combine with each other.

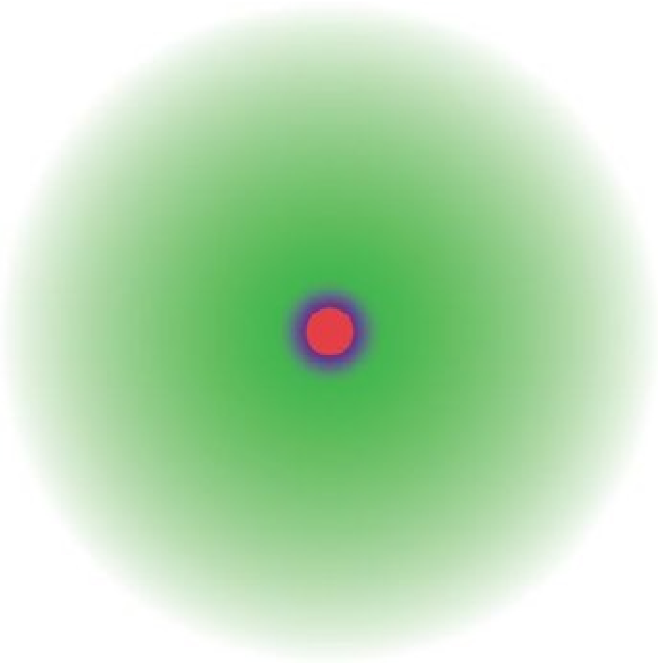


Fig. 4-2. The same atomic nucleus shown in Fig. 3-1(a), with its surrounding “green” EM field, but with a purple halo added to represent the short-range strong field that extends a short distance outside the nucleus (not to scale).

Yukawa’s quantum. Yukawa realized that, just as EM radiation is made of discrete quanta called photons, so his strong field could exist as separate quanta. And just as photons are described by Maxwell’s equations, so the propagation of these quanta would be described by the field equation he had derived. This is where Yukawa came up with his second surprise: The λ term that he added to give the field a short range also affects the propagation of a quantum, causing resistance to changes.[5]

A field with mass. In classical physics, mass or inertia is the tendency of an object to resist a change in its motion (see Chapter 2). Thus, while photons and gravitons travel at the speed of light, the λ term would cause Yukawa’s quanta to propagate through space more slowly — in fact, like a particle with mass. When Yukawa calculated the effective mass of his field from the experimentally-determined λ, he found a value about 200 times larger than the electron mass, but less than the proton mass. Because of this, the strong field quantum was given the name *mesotron* (from *meso* meaning intermediate), later shortened to *meson*.

> **Until 1935… it was thought that only particles with zero mass, such as photons, could be the quanta of a force field. Yukawa**

> **showed that this is not so and that particles (now called… mesons) can also be the quanta of a force field. That Yukawa should have arrived at this important discovery before any of the European and American physicists did is all the more remarkable when one realizes that… Yukawa had very little contact with European physicists. – H. Boorse & L. Motz (B1966, p. 1417)**

A field with charge. Mass was not the only surprise that Yukawa found. In order to exert a binding force between protons that are positively charged and neutrons that carry no electric charge, the new field must itself carry an electric charge. In fact, there must be two charged fields – one positive and one negative. (A third field with no charge was added later.) Up to this point, fields were properties of space whose only physical manifestation is that they exert a force, but now physicists had to cope with a field that has both mass and charge — properties previously associated only with particles.

To those learning physics for the first time, it may not be a surprise to learn that fields can have mass and charge. After all, if *everything* is made of fields, then all physical properties – spin, mass, charge, energy, etc. – must be properties of fields. But those with prior exposure to physics have come to see these properties as belonging to "solid" matter, so picturing them as properties of an intangible field may not be easy. All we can do is point to a term in the field equation that slows down propagation of the field and hence corresponds to mass, and point to an interaction term that describes how the Yukawa field interacts with the EM field, giving it an effective electric charge. It is these mathematical terms in the field equations that give the field the mass and charge properties that we ordinarily ascribe to "material" objects.

A field without spin. Although the strong field has mass and charge, in another respect it is simpler than the EM field. You recall that quantum fields have a property called spin, or helicity, that is related to the internal complexity, and that the EM field, with electrical and magnetic components that can point in various directions, has a spin of 1. Well, the strong field does not have the same complexity; its force is always attractive and its spin is zero. Now I can hear some of you screaming, "Wait a minute! The gravitational field also consists of a single attractive force, and yet you said its internal complexity is greater and its spin is 2." This is true. For reasons that are beyond the scope of this book, Einstein's gravitational field is more complex than any other field. I must ask you to take it on faith that the strong field is simpler than both the gravitational field and the EM field, and that its spin is 0.

Field or particle. Yukawa's field with mass and charge further narrowed the difference between fields and particles. After all, if a field quantum is emitted as a unit from a single location (like a particle), is absorbed as a unit at another location (like a particle), and in between exhibits a mass and an electric charge (like a particle), how is it different from a particle? There is no empirical test that can answer this. The difference lies in how we choose to perceive reality – the picture in our mind, and this difference is immense. A particle is a localized bounded object. It has edges, or perhaps, as some think, it is only a point. A particle exists *in* space and around it is empty space. A field, on the other hand, is something that exists everywhere as a property *of* space. Its intensity may be small, but it is never zero. Even in areas where there are no quanta there is a small amount of field called the vacuum field. In QFT there is no such thing as empty space. While particles move through space like baseballs, fields are diffuse and spread-out; they exhibit interference and diffraction. Nevertheless, when a spread-out quantum is finally absorbed, it collapses into one location (see quantum collapse in Chap. 10).

When Yukawa's theory came along, the particle picture was strongly embedded and QFT was in its infancy. Even the long-held view of light as a field was doubted, with most physicists following Einstein in his belief (at the time) that photons are particles. And so Yukawa's meson was seen by most physicists as a particle. However, Yukawa himself was careful to use the term quantum – the term Planck introduced to denote a discrete amount of field. While I have been using the word particle in accord with custom, please understand that these particles are really fields — or, I should say, quanta of fields. *In QFT* th*ere is no such thing as a particle*.

Where is it? Whether it is a quantum of field or a particle, Yukawa had predicted a new entity – something not previously seen or even dreamt of. Why, wondered Yukawa, had it not yet been seen?

> **As such a quantum with large mass and positive or negative charge has never been found by experiment, the above theory seems to be on a wrong line… The reason why such massive quanta, if they ever exist, are not yet discovered may be ascribed to the fact that the mass is so large that [the necessary energy relationship] is not fulfilled in ordinary nuclear transformations. – H. Yukawa (B1966, p. 1423-1425)**

In other words, Yukawa's quanta had not been observed because there wasn't enough energy to create one. While photons can be created with very little energy — for example, by heating up atoms — to create a field quantum with a mass 200 times that of an electron would take more energy than is found in an ordinary atom. Even radioactive nuclei do not have enough energy to create such a massive particle. This led Yukawa to make a prediction that in its audacity rivaled Einstein's prediction about the bending of light:

> **These massive quanta may also have some bearing on the showers produced by cosmic rays. – H. Yukawa (B1966, p. 1427)**

Like Einstein's famous prediction, Yukawa's was eventually verified, but the path to verification was not as quick or direct, and it involved one of the most famous "false positives" in physics history. To understand the story, we must first understand cosmic rays.

Cosmic rays. Cosmic radiation was discovered in 1912 by the Austrian physicist Victor Hess. Like many discoveries, it was a bit of an accident. At the time, natural radioactivity was known to exist at the earth's surface, coming from minerals like uranium, radium, etc. Hess therefore began a series of balloon ascents in the hope of measuring the falloff in radiation at higher altitudes. Much to his surprise, he found instead that the radiation began to *increase* at 1 km above sea level; in fact, at 5 km it was several times greater than at sea level. He concluded that this radiation could only come from outer space, hence the name *cosmic rays*. Hess even made an ascent during a solar eclipse to see if the sun was the source. (It wasn't.) The source is still not known, but we do know that cosmic rays consist mostly of atomic nuclei with very high energy that sometimes disintegrate and produce the "showers" that Yukawa referred to. Cosmic rays are believed to be responsible for Darwinian mutation, computer crashes, and cancer, and for a while they served as the physicists' "nuclear physics laboratory".

False validation. It only took two years for Yukawa's prediction to be verified (less than for Einstein's prediction) — or at least, so it seemed at the time. In 1936 a particle was observed in cosmic rays with a mass close to Yukawa's prediction, and with electric charge. The excitement in the physics community was immense, but the bubble burst when it was discovered that the new particle did not interact appreciably with nucleons — the *sine qua non* for a nuclear force field. This caused Isidor Rabi, 1944 Nobel laureate, to exclaim "Who ordered that?".

True validation. The dilemma was resolved ten years later, thanks to an improvement in detecting cosmic rays made by Cecil Powell, an English physicist. Powell's idea was to record radiation tracks in a photographic emulsion, rather than in the previously-used cloud chambers.

> **[The method] has the advantage of providing a remarkably extended time-scale for the study of certain types of transient phenomena. In particular, particles of very short life-time… which commonly decay 'in flight' when moving in the atmosphere, are arrested in a solid material in a time interval more than a thousand times shorter than when moving in a gas. It is thus possible to study the spontaneous decay of the particles when stopped in an emulsion, or their interactions with nuclei, and thus to observe phenomena which it has proved very difficult or impossible to record by other methods. – C. Powell (B1966, p. 1435-6)**

Powell collected data from a number of high altitude sites, including the French Pyrenees and the Bolivian Andes. In 1947 he found tracks made by a particle with a mass not much different from the meson discovered in 1936, but that interacted with nucleons as Yukawa had predicted. Yukawa was awarded the 1949 Nobel Prize and Powell's followed a year later.:

> **Professor Hideki Yukawa. In 1934, when you were only 27 years old, you boldly predicted the existence of new particles, now called "mesons", which you anticipated to be of fundamental importance for the understanding of the forces acting in the atomic nucleus. Recent experiments have provided brilliant support for your essential ideas… you have played a great role in bringing your country to its very high position in modern physical research. – I. Waller (Nobel presentation speech, 1949)**

Two mesons? When Einstein was asked why he didn't use shaving cream, he replied, "Two soaps? That's too complicated." Physicists had similar problems with the two mesons. To minimize confusion, Powell added Greek letter prefixes, calling the 1936 particle a mu-meson and his 1947 particle a pi-meson, later shortened to *muon* and *pion*. When the pion was recognized as the Yukawa meson, the muon became a kind of orphan child; it didn't fit any theory. The story eventually had a happy ending when the muon found its place in QFT as a kind of "fat sibling" of the electron.

Yukawa's error. Despite Yukawa's great success, he got one detail wrong. His mistake had to do with the mathematical difference between *scalar* and *pseudoscalar* fields. Even to define these terms would take us beyond the scope of this book, but the difference was serious enough to create discrepancies between his theory and later experimental results. The difficulty was resolved in 1941 by Julian Schwinger and William Rarita, who showed that the strong field is, in fact, pseudoscalar. This paper was published when Schwinger was 23 years old.

Quarks. I would like very much to leave the subject at this point, with the Yukawa field and its pions proving a great success. However, a major change in our understanding of the fundamental fields occurred in the 1960's, as more quanta of the strong field and the nucleon field were discovered – in fact, an embarrassingly large number. This led to a variation of QFT called *quantum chromodynamics,* in which the strong field is composed of more basic fields called quarks and gluons, even though these basic fields are never seen (see "The invisible quark" in Chapter 6). But don't worry; this need not concern any but physicists. Whether it is a basic field or a composite of quark and gluon fields, the behavior of the strong field remains the same.

SUMMARY

The discovery of the atomic nucleus led to the discovery of a new force field that binds protons and neutrons tightly together in the atomic nucleus. It is called the strong field and its quanta are called pions. The strong field has the following characteristics:

- It is generated by and acts on nucleons.
- Its force is always attractive, like gravity.
- It is strong enough to overcome the electric repulsion between protons (hence its name).
- It has a short range, with an intensity that drops almost to zero at distances greater than the size of a nucleus.
- It has a mass that causes pions to propagate more slowly through space than photons. The mass and short range arise from the same term in the field equations.
- It comes in three electrical varieties: positive, negative and neutral.
- It has zero spin.
- Because of the mass and charge, the field quanta behave more like particles than do photons.

The three fields presented so far can be summarized as follows:

Field	Spin	Quantum	Interacts with	"Color"
Gravity	2	Graviton	everything	blue
EM	1	Photon (γ)	charged fields	green
Strong*	0	Pion (π^+, π^-, $\pi°$)	Nucleons	purple

*We will see in Chapter 10 that the strong field is made of more basic but "invisible" fields called quarks and gluons.

CHAPTER 5

THE WEAK FORCE - The "brown" field

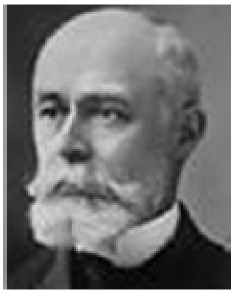
> **An old dream was fulfilled last year when the discoveries of the W and Z were made at CERN[1] — the dream of better understanding the weak interaction… The weak interaction is unique in that it can change the nature of a particle, for example transforming a neutron into a proton or vice versa. – G. Ekspong (Nobel award speech, 1984)**

The dream was quite old indeed. It began in 1896 with the discovery of *radioactivity*: the process in which atomic nuclei spontaneously emit particles or rays, and it wasn't completely fulfilled until the experimental confirmation of the weak field theory in 1995. During those 99 years there were enough twists, wrong turns and experimental errors to fill a volume. More than a dozen Nobel laureates played a part, new particles had to be discovered, and a fundamental law had to be overthrown before the struggle was over. In the end, the *weak field* emerged and took its place alongside the gravitational, EM, and strong force fields in the QFT family. Like the strong field, the weak field has both mass and charge. Reflecting the tortuous path travelled before the final understanding was reached, the weak field is represented by two symbols: W for the charged fields and Z for the neutral field.

Before beginning the story, I want to warn you that the going gets rough. Even though I avoid mathematics and stay at a relatively superficial level, there are concepts that some may find hard to digest. These people may safely skip this chapter without losing the basic appreciation and understanding of quantum fields. The weak field, after all, is not as important to this understanding as the matter fields that will be presented in the next chapter. However, if you do continue, you will find that the story of the weak field is a most fascinating one, and I think you will be well rewarded for your patience and persistence.

RADIOACTIVITY

The discovery of radioactivity, like so many discoveries in physics, was accidental. It happened because a French physicist couldn't wait for the sun to shine.



Antoine Becquerel (1852-1908). Antoine Becquerel was a professor of physics at the Paris Museum of Natural History, following in the footsteps of his father and grandfather before him. The great achievement of his life occurred in 1896 when he learned of the exciting discovery of x-rays by Wilhelm Roentgen. Becquerel's father had worked with phosphorescent materials (substances that give off a glow after being exposed to light), so it was natural for Antoine to wonder if these chemicals might also emit the newly-discovered x-rays.

> **"On the very day that news reached Paris of the experiments of Roentgen and of the extraordinary properties of the rays emitted by the phosphorescent walls of Crookes' tubes,[2] I thought of carrying out research to see whether all phosphorescent material emitted similar rays. The results of the experiment did not justify this idea, but in this research I encountered an unexpected phenomenon." – A. Becquerel (Nobel lecture, 1903)**

Becquerel's experiment was to wrap a photographic plate in black paper and place it under one of his father's chemicals which, as it happened, contained uranium. He then left it in the sunlight, hoping that the phosphorescent chemical would emit Roentgen's x-rays that would pass through the black paper and expose the photographic plate. Eureka! When he developed the plate it was indeed fogged, and so he concluded that the chemical had been stimulated by the sunlight to emit x-rays. But then a funny thing happened.

After several days of intensive study with other plates, there came a succession of cloudy days. Becquerel grew impatient. He decided to develop the plate anyway, thinking he might see slight fogging. To his surprise, he found that the fogging was just as intense as when there was sunlight. After experimenting with other chemicals, he determined by the process of elimination that the emissions came from the uranium, and that they were constant and continual, unaffected by sunlight or other factors. He had discovered *radioactivity*!

Marie (1867-1934) and Pierre (1852-1908) Curie. Shortly after this discovery, Becquerel suggested to a student at the Sorbonne that she should look for similar emissions from other chemicals. The student was Marie Curie, newly married to a physics professor named Pierre Curie. Recognizing the importance of the work, her husband joined in her research, and between them they soon identified other radioactive elements, including two not previously known that Marie named radium (because it radiates) and polonium (after her native Poland). It was Marie Curie who coined the term *radioactivity* to describe the mysterious process. The discovery of radium was especially important because its radiation is about a million times stronger than that from uranium.

Enter Rutherford. The following year, the 26-year-old Ernest Rutherford, fresh from his B.A. degree from Trinity College and with the discovery of the nucleus still ahead, decided to take on these mysterious radiations. In typical Rutherford fashion, he ended up “writing the book”. He recognized the random nature of the radiations and was the first to use *half-life* as a measure of the decay rate.[3] He classified radioactivity into three types: *alpha*, *beta* and *gamma* — names that persist to this day.[4] Fourteen years later, Rutherford showed that all three radiations emanate from the atomic nucleus, and in the case of alpha and beta radiation the nucleus ends up with a different atomic number. In other words, it becomes a different chemical element, which is why the word *decay* is applied to the process. It is *beta decay* that we are concerned with now, because that is what led to the discovery of the weak field.

Nobel Prizes galore. Half of the 1903 Nobel Prize was given to the two

Curies, with Becquerel receiving the other half. Rutherford received the 1908 Nobel Prize in chemistry for “investigations into the disintegration of the elements, and the chemistry of radioactive substances”. Like any good physicist, he was a bit chagrined at his “demotion” to chemist.

> **Rutherford declared that he had dealt with many kinds of radioactive transformations with different periods of time, but that the quickest transformation he had met was his own transformation from a physicist to a chemist in a single moment — (B1966, p. 703).**[5]

Beta decay. The next step was made in 1900 by Becquerel. He showed that Rutherford’s beta rays are really electrons, which had been discovered in 1897 by J.J. Thomson. It was also discovered that when a nucleus emits beta radiation, it ends up with one more proton and one less neutron than when it started. Somehow a neutron was emitting an electron and changing into a proton. This is possible mass-wise because the neutron is heavier than a proton and electron combined, so mass was not being created. It also made sense in terms of nuclear stability, since beta decay occurs only in nuclei that are neutron-rich to begin with (see Chap. 4, “A delicate balance”), so changing a neutron into a proton would bring the number of neutrons closer to the optimum value. But this doesn’t say *how* beta emission occurs. What was needed was a mechanism, and at the time there was no force that could explain why a neutron would change into a proton. Beta decay remained a mystery.

Carbon dating. Before telling you the answer to the mystery, I’d like to digress for a moment and tell you about an important application of beta decay known as *carbon dating*. This is a method of determining the age of organic materials by measuring the concentration of C-14, an isotope of carbon. While C-14 is chemically identical to ordinary carbon, it is radioactive and decays into nitrogen (N-14) with a half-life of almost 6,000 years. In the atmosphere there is only one C-14 atom for every trillion carbon atoms and if not for the radioactive emissions we would never know it is there. This ratio of one trillionth remains steady because the amount lost by radioactive decay is replenished from the upper atmosphere, where nitrogen is continually being converted into C-14 by cosmic rays. The same ratio (one trillionth) is found in living plants and animals because their source of carbon, directly or indirectly, is the atmosphere. However, after an organism dies there is no more replenishment from the atmosphere — only decay — and so the amount of C-14 begins to decrease slowly. By comparing the amount of beta radiation from a once-living substance (wood, paper, an Egyptian mummy, etc.) with the total amount of carbon, one can determine the C-14 ratio and infer how long the substance has been dead. For discovering this ingenious method, Willard F. Libby was awarded the 1960 Nobel Prize in chemistry.

> **Seldom has a single discovery in chemistry had such an impact on the thinking of so many fields of human endeavour. Seldom has a single discovery generated such wide public interest. – A. Westgren (Nobel award speech, 1960)**

Now let us return to our story and see how the beta decay quandary led to a theory of the weak field. It is a long and complicated story, and it began with the prediction of a new particle.

THE NEUTRINO

In 1930, four years before Yukawa introduced the strong force field, there were only two known force fields — gravity and EM, and three known atomic particles—the electron, proton and neutron. (I am still calling them particles as per conventional usage, but please remember that in QFT they are field quanta.) To add another particle or field to this simple, satisfying picture of nature seemed unthinkable. On the other hand, there was this nagging mystery of beta decay.

Pauli's postulated particle. A clue to the mystery was that the emitted beta rays have different energies, indicating that some energy is going elsewhere. This led Wolfgang Pauli to suggest that a second particle is emitted from the nucleus at the same time as the electron, carrying with it the unaccounted-for energy. He dubbed his hypothetical particle the "neutron", but when that name was usurped by Rutherford two years later, Enrico Fermi renamed Pauli's particle the *neutrino* (Italian for "little neutral one").

> **"I admit that my expedient may seem rather improbable from the first, because if neutrons [i.e., neutrinos] existed, they would have been discovered long since. Nevertheless, nothing ventured, nothing gained… We should therefore be seriously discussing every path to salvation." – W. Pauli (quoted in F. Reines' Nobel lecture, 1995)**

Because of its very weak interaction with matter, the neutrino remained elusive. "I have done a terrible thing," moaned Pauli, "I have postulated a particle that cannot be detected". Hans Bethe agreed, saying "there is no practically possible way of observing the neutrino." (When teased about this after the discovery was made, Bethe said, "Well, you shouldn't believe everything you read in the papers.") Nevertheless, despite the lack of experimental evidence, the neutrino was soon accepted on theoretical grounds as an essential part of beta decay.

Confirmation. It would be 26 years before Pauli's postulated particle (say *that* three times fast!) was observed in a heroic experiment by Frederick Reines and Clyde Cowan, who took on the project because "everybody said you couldn't do it". Because its interaction with matter is so weak, they needed a copious source of neutrinos and being at the Los Alamos Laboratory, their first thought was to use a nuclear bomb test, but the experiment was actually performed at a nuclear power plant on the Savannah River in South Carolina. When Pauli was notified of their success by telegram, 26 years after his prediction, he wired back, “Thanks for message. Everything comes to him who knows how to wait.” Half of the 1995 Nobel Prize in physics was awarded to Reines (Cowan having died in 1974). Reines, who had to wait 39 years for the award, could have given Pauli a lesson in waiting!

All in the family. Quantum fields come in families — fields that have the same spin and obey similar equations. We have seen that the EM field contains electric and magnetic fields, and the strong field is a family of three fields with charges +1, -1, and 0. We also learned in Chapter 4 that the muon is a “fat sister” of the electron, so it should come as no surprise that the neutrino found a place in that same family, as a “baby sister” to the electron and muon. These three particles (electron, muon and neutrino) are called *leptons* (from Greek “leptos”, meaning small), while the heavier nucleons (proton and neutron) are called *baryons* (from the Greek “barys” meaning heavy). However, it is not the neutrino that we are concerned with here. Its discovery was just another step leading to the field that causes beta decay.

THE WEAK FORCE FIELD

Enrico Fermi (1901-1954). It didn't take long after Pauli postulated the neutrino before someone came up with a theory to describe the process in which it was emitted. That someone was Enrico Fermi, four years after he gave the neutrino its name. Just as a photon is emitted when an atomic electron changes its energy state, so Fermi's equation showed how a neutron can change its energy state and become a proton while emitting an electron and a neutrino. Now there are several ways this equation could be constructed and, as it happened, Fermi chose the wrong one, but that didn't become apparent until 20 years later. (The same thing had happened with Yukawa's strong field equation which, as we saw in Chapter 4, was corrected seven years after it was introduced.)

A more serious problem was that Fermi's equation lacked a field. The first person who tried to fill this gap was Yukawa. In the same paper in which he derived equations for the strong spin-0 field, he suggested that this same field could explain the Fermi equation. As it turned out, this idea didn't work because 0 was the wrong spin, but Yukawa was right about two things: the field responsible for beta decay must have mass (to give it a short range) and charge (so that it can convert a neutron into a proton). It took the son of a Swedish rabbi to get the spin right.

Oskar Klein (1894-1977). Oskar Klein was an under-appreciated Swedish physicist who is not even listed among the 1500 scientists in "Asimov's Biographical Encyclopedia of Science and Technology" (A1982). Klein realized that the field needed to explain beta decay must have spin 1, like the photon, not 0 as Yukawa thought.[6] Klein presented his theory at the 1938 Warsaw Conference, but his contribution was largely overlooked until being revived (or rediscovered) by Schwinger in 1957.

Color it brown. We need a color to help us visualize the weak field. I've run out of cool colors (green, blue, purple) for force fields, so I will choose brown. Brown is rather colorless and thus appropriate for a field that is so weak. Since the weak field interacts with all matter, you should picture a faint brown halo around every nucleon, electron, neutrino, etc. The halo is even smaller and tighter than the "purple" halo around nucleons, because the weak field has a shorter range, its mass being larger by a factor of 500.

How fields cause decay. The idea that a field can cause a particle to decay, i.e., transform itself into other particles, was a new one. The process consists of two phases. The first phase is similar to the way an electron emits a photon while dropping to a lower energy state in an atom. In the same way, the field equations for the weak force show that a neutron can change to a lower energy state (i.e., become a proton) while emitting a quantum of the weak field, except that the mass of the weak field quantum is so large that there is not enough energy to create a fully independent quantum. Instead what is created is a kind of *incipient quantum* that is intermediate between a self-field and a separate quantum. (Particle people call it a *virtual* particle.) While the incipient quantum doesn't have enough energy to fly off and become independent, there is enough energy to create an electron and a neutrino. Thus beta decay occurs when a neutron changes into a proton while emitting an incipient "brown" weak field quantum, which then transforms into an electron and a neutrino. The quantum of the weak field has been given the name *weakon.* (I would have preferred *adynaton*, from the Greek word for weak, but I didn't get to vote.)

It is tempting to end the story at this point. After all, the essence of the weak field has now been presented and the rest of the story is, in a sense, nothing more than getting the algebra right. If you found the going rough so far, you may skip this algebra and go directly to the summary.

THE RIGHT EQUATION

Fermi's equation for beta decay was quickly accepted, but as more data were collected it became apparent that the equation didn't quite work. Attempts to find the right equation were foiled by a fundamental law called *parity conservation*.

Parity. Parity conservation is the physicist's way of saying that nature's laws are left-right symmetric – that there is no preference for left or right. This statement is usually illustrated by the *mirror test*.[7] Imagine that you are looking at some process in a mirror. If nature is left-right symmetric, then whatever you see in the mirror can also happen in real life. This law had been unquestioned for centuries, but doubts began to arise at the 1956 Rochester Conference on High-Energy Physics.

> **I was sharing a room with a guy named Martin Block, an experimenter. And one evening he said to me, 'Why are you guys so insistent on this parity rule?' I thought a minute and said, 'It would mean that nature's laws are different for the right hand and the left hand'. The next day at the meeting, Oppenheimer said, 'We need to hear some new, wilder ideas about this problem.' So I got up and said, 'I'm asking this question for Martin Block: What would be the consequences if the parity rule was wrong?'… Lee, of Lee and Yang, answered something complicated and as usual I didn't understand very well. – R. Feynman (F1985, p. 247-250)**

Tsung-Dao Lee (1926-) and Chen Ning Yang (1922-). T.D. Lee (left) at Columbia University and C.N. Yang at the Princeton Institute for Advanced Studies were the first to point out that there is no theoretical or experimental basis for the belief in left-right symmetry. They also suggested an experiment to test it. The experiment is not hard to understand; all you have to know is that nucleons, like photons, have spin (see Chapter 6). This spin causes the nucleon to become a tiny magnet in the same way that spinning electrons create the magnetism in ordinary magnets (Chapter 3).[8] It follows that if a nucleon is placed in an external magnetic field, it will tend to align itself along the field, just as a compass needle aligns itself along the earth's magnetic field. Yang and Lee's idea was to take a radioactive substance that emits beta rays and see if the number of beta rays emitted in a given direction remains the same when the nuclear spins reverse their direction (Fig. 5-1).

Fig. 5-1. The mirror test applied to the Wu experiment. The fingers (on the left) represent the rotation of nuclear spins while the thumb indicates beta rays being emitted in the upward direction. In the mirror image (on the right) the direction of rotation indicated by the fingers is reversed, but the thumb still points upward. If nature is left-right symmetric, the beta radiation in the upward direction should be the same either way.

Chien Shiung Wu (1912-1997). Lee and Yang's proposed experiment was carried out at Columbia University by a team under the direction of another Chinese-American, Chien Shiung Wu, using cobalt-60 as the beta-emitter (see Fig. 5-2). An electric coil was used to produce a magnetic field to align the nuclear spins and the sample was cooled almost to absolute zero to enhance the alignment. After the number of beta rays emitted upward was counted, the magnetic field was reversed to reverse the spin direction and the radiation was counted again. If parity is conserved, the two counts should be the same.

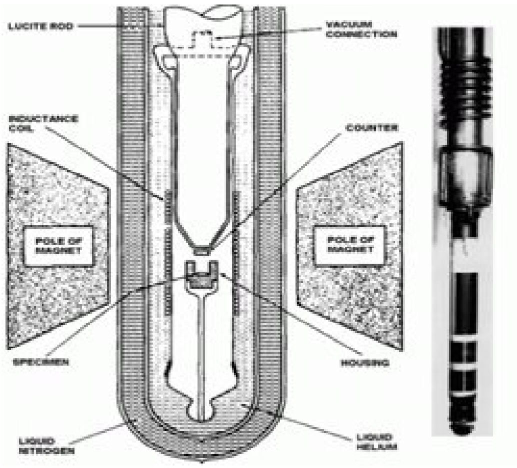


Fig. 5-2. Apparatus for testing parity conservation (diagram at left, photo at right). The magnet poles were used to lower the temperature of the specimen almost to absolute zero by a process called *adiabatic demagnetization*. The nuclei in the specimen were then aligned using the magnetic field produced by a surrounding electric coil (not shown) and the beta radiation was measured by the counter. The alignment was then reversed by reversing the current in the coil and the counting repeated. According to the mirror test (Fig. 5-1), the result should be the same. (www.physics.nist.gov)

Get your bets in! Needless to say, interest was high in Wu's experiment. In fact the physics community began to resemble a pari-mutuel window at a race track. Yang and Abraham Pais each bet a dollar that the result would be left-right symmetric. Wolfgang Pauli went further, saying "I am ready to bet a very large sum that the experiments will give symmetric results" (G1992, p. 334), and Feynman was willing to lay 50-to-1 odds on symmetry. (Later he said he was proud that he offered *only* 50-to-1.) Schwinger refused even to bet, saying "I could not accept that nature could be so mischievous as to destroy one of the symmetries" (M2000, p. 417).

And the winner is… To everyone's surprise, the number of beta rays emitted upward did change with the direction of spin, thereby violating the mirror test. Nature, it seems, is left-right *asymmetric*, at least when it comes to beta decay, and everyone lost their bet. However, Lee and Yang won the 1957 Nobel Prize in physics while, to the disappointment of many, Dr. Wu was not included in the award. With parity overthrown, the way was now clear for replacing Fermi's incorrect equation (called *T*) with the correct equation (called *V and A*).[9] However, there was a problem with the large amount of data that supported the *T*-equation.[10] It was only when these data were shown to be incorrect that further progress could be made.

<u>Feynman's role</u>. At the next Rochester Conference in 1957, Lee presented a paper saying that while most experimental data on beta decay supported the *T* equation of Fermi, the Wu result suggested that it should be *V and A*. One of the attendees at the meeting was Richard Feynman. After studying a copy of the talk, Feynman realized that some earlier work he had done on beta decay might be relevant after all, except for the experimental data. After returning from a summer in Brazil playing bongo drums (among other things), he learned that people were beginning to be suspicious of those data.

> **When I got back to Caltech, I asked some of the experimenters what the situation was with beta decay… "The situation is so mixed up [they said] that even some of the things they've established for years are being questioned – such as the beta decay of the neutron is… T. It's so messed up. Murray [Gell-Mann] says it might even be V and A." I jump up from the stool and say, 'Then I understand EVVVVVERYTHING!' They thought I was joking. But the thing that I had trouble with at the Rochester meeting – the neutron and proton disintegration; everything fit but that, and if it was V and A, that would fit too… That night I calculated all kinds of things with this theory… I went on and checked some other things, which fit, and new things fit, and I was very excited. It was the first time, and the only time, in my career that I knew a law of nature that nobody else knew. – R. Feynman (F1985, p. 247-250)**

<u>Gell-Mann's role</u>. Murray Gell-Mann had found the *V and A* equation before Feynman, and he wasn't happy to hear Feynman carrying on about his great discovery "that nobody else knew". Gell-Mann was Feynman's colleague and one-time friend at the California Institute of Technology. The pair were a true "odd couple", with Feynman playing Oscar Madison to Gell-Mann's Felix Unger. Sidney Coleman put it this way: "Murray's mask was a man of great culture. Dick's mask was… a little boy from the country that could see through things the city slickers can't" (G1992, p. 189).

> **The real trouble started… when Gell-Mann returned from a wilderness trek in northern California to find Feynman strutting around the department proclaiming an epiphany: in a moment of intense clarity he had seen the very secret of what is known as the weak nuclear force, the engine of radioactive decay. Gell-Mann**

was dumbfounded. He had been mulling over the problem for weeks, and had come to a similar conclusion. But, typically, he hadn't written the idea up… Now fearful of being scooped, he rushed to beat Feynman into print. Finally the department chairman intervened, persuading his two stars that it would look silly to have competing papers come from the same campus. They agreed to collaborate… By the mid-1980's what had begun as a friendly rivalry was a bitter feud… The breaking point came in 1985… [when] Gell-Mann found Feynman's account [F1985] of the theory of the weak nuclear force, the one they had reluctantly collaborated on… Gell-Mann, enraged, said he would sue. In a later edition Feynman conceded that Gell-Mann and two other physicists had also thought of the idea. But the disclaimer didn't heal the wound. — G. Johnson ("The jaguar and the fox", Atlantic Monthly, July, 2000)

<u>**Schwinger's role**</u>. Another physicist who found the *V and A* equation before Feynman, and also before Gell-Mann, was Julian Schwinger, the "gentle genius of physics" (M2000, p. 424). Starting from the assumption that the weak force is transmitted by a field with charge and mass (like the strong field), but with spin 1 (like the EM field), he postulated two charged fields that he called Z^+ and Z^-. From that assumption he derived the *V and A* equation as the most natural form for the weak field (S1957, p. 424), but he too had to face the large amount of data at the time that pointed to a *T* equation. This led him to suggest, among other possibilities, that the weak field behaves phenomenologically like a *T* equation, even though the underlying equation is *V and A*. Of course when the experimental data were corrected there was no longer need for such a band-aid. It is understandable, therefore, that Schwinger was also unhappy with Feynman's claim:

It has given me no pleasure because my prior discovery has never been recognized. Isn't that funny? Well, too many people had too much at stake to say, 'Oh yeah, but…' I think I'm angry still. Do I have a right to be?… I think Feynman gets credit… [when] experiments were beginning to get shaky – for insisting that maybe the experiments are wrong. But at this stage there was no possibility of questioning the experiments. I had no idea that the experiments were so tricky and unreliable. Who would? So my fault was not simply to believe enough in what I had done and to try desperately to accommodate the experiments. Yet every time I tried to accommodate, it didn't work. – J. Schwinger (M2000, p.

426)

The different attitudes of Schwinger and Feynman toward experimental data say much about their different personalities. Compare Schwinger's phlegmatic and poignant "Who would?" with Feynman's more aggressive militant attitude:

> **When I became interested in beta decay directly, I read all these reports by the 'beta-decay experts', which said it's T. I never looked at the original data; I only read those reports, like a dope. Had I been a good physicist, when I thought of the original idea back at the Rochester Conference I would have immediately looked up 'how strongly do we know it's T?' – that would have been the sensible thing to do… Since then I never pay any attention to anything by 'experts'. I calculate everything myself… I'll never make that mistake again… Of course, you only live one life, and you make all your mistakes, and learn what not to do, and that's the end of you. – R. Feynman (F1985, p. 254-255)**

<u>Sudarshan's role</u>. Schwinger, Gell-Mann and Feynman were not the only ones to solve the conundrum. George Sudarshan, a young physicist working under Robert Marshak at Rochester University, also had discovered the *V and A* equation. He wanted to present his results at the same 1957 Rochester conference where Lee raised the *V and A* question, but couldn't because he was only a student.

ELECTROWEAK UNIFICATION

Now that the correct equation was found, the task was to find the field responsible for the process and place it in the right family. Since the field must have spin 1, you will not be surprised to learn that it wound up in the same family as the EM field, the only other field with spin 1. But once again, working out the details wasn't easy, and once again it was Schwinger who made the first step, in the same paper where he introduced the *V and A* equation. Using the lepton family (electron, muon and neutrino) as a model, Schwinger suggested that the two weak fields that he had postulated (Z^+ and Z^-) could be joined with the neutral EM field to make a family of three fields with spin 1.

> **From the general suggestions of a family of bosons[11] that is the analog of the leptons, and the identification of its neutral member as the photon, we have been led to a dynamics of a charged, unit-spin Z-particle field that is interpreted as the invisible instrument of the whole class of weak interactions. – J. Schwinger (S1957, p. 433)**

In postulating only two charged weak fields, Schwinger made the same mistake that Yukawa had made about the strong field. It was Schwinger's student, Sheldon Glashow, who added a neutral weak field. Ironically, Schwinger's Z notation survived for the neutral field that he did *not* introduce, while the ones he *did* introduce were later renamed W.

Sheldon Glashow (1932-). As Schwinger's doctoral student, Sheldon Glashow was given the task of developing Schwinger's idea that the weak field was part of a family of three fields with spin 1.

> **Julian was convinced of the existence of an 'intermediate vector boson' and of a fundamental connection between weak interactions and electromagnetism… My task was not precisely delineated. It was to seek and perhaps find such a relation, and to explore its**

observable consequences… In those days of yore, our understanding of the microworld was expanding at breakneck speed. A once theoretically 'dictated' and experimentally 'established' parity-conserving model of the weak force was bit by bit giving way to the correct parity-violating V-A picture… He convinced himself (and me!) that a triplet of vector bosons … could possibly offer a plausible, elegant, and unified explanation of all EM and weak phenomena. – S. Glashow (N1996, p. 157-158)

After finishing his thesis, Glashow continued his work at the Bohr Institute in Copenhagen. It was there that he realized that if weak interactions violate parity conservation while EM interactions do not, they cannot be as closely related as Schwinger thought. This led him to add a neutral weak field that he called Z°, following Schwinger's Z-notation, while moving the photon to a more "cousinly" relationship. "It took me over a year to see this, since I no longer had direct access to Julian," he said later (N1996, p. 159). His article was published in 1961.

Confirmation. Just as with Pauli's neutrino, it was clear that detecting this quantum (or "particle") would pose a serious challenge:

The direct identification of this hypothetical particle will not be easy. Its linear couplings are neither so strong that it would be produced copiously, nor are they so weak that an appreciable lifetime would be anticipated – J. Schwinger (S1957)

In 1983 evidence for the weak field quantum was finally obtained at the giant CERN accelerator in Geneva, Switzerland. In fact, all three quanta were detected: Schwinger's charged fields and the neutral field that Glashow introduced. For this achievement, Carlos Rubbia and Simon van der Meer were awarded the 1984 Nobel Prize in physics.

However, there was still a remaining problem. The mass of the new field turned out to be 500 times greater than that of the strong field, making it the heaviest known quantum field, and its range therefore the shortest. Even before the weakon was detected, it was known that its mass had to be very large to explain the feebleness of the weak interactions. An explanation for this large mass had to be found.

THE HIGGS FIELD

The first person to find the explanation, once again, was Julian Schwinger, in the same paper (S1957) in which he introduced the *V and A* equation. He showed that if there is a certain kind of field spread throughout space, quanta moving through it would be slowed down, just as mass slows the speed at which fields propagate. As Schwinger said, "**My idea here was, from the very beginning, to use the scalar field as a way of generating masses**" (M2000, p. 421). As usual, Schwinger's contribution was overlooked and forgotten. A similar mechanism was proposed in 1964 by Robert Brout, Peter Higgs and François Englert, and the field is known as the Higgs field.

The Higgs boson. Before the Higgs field could be accepted, it was necessary to identify its quanta, now called *Higgs bosons*, and this would not be easy. It wasn't until July 4, 2012, that a particle believed to be the Higgs boson was detected at the CERN laboratory. François Englert and Peter Higgs were then awarded the 2013 Nobel prize in physics (Brout having died before the award was given).

The Higgs mechanism was incorporated into the weak field equations by Steven Weinberg and Abdus Salam, working independently. In the process Weinberg changed Schwinger's Z notation for the charged weak fields to W (for weak, or possibly Weinberg :<), but he retained Z for the neutral field, resulting in the present hybrid notation. For their achievements, Glashow, Weinberg and Salam shared the 1979 Nobel Prize, while Schwinger's contributions were, as usual, forgotten.

A final fix. At this point, you must be hoping that the long saga is over. Enough already with *T, V and A,* parity, Higgs mechanisms, and fields with mixed up names. But there was still one remaining hurdle: there were calculational difficulties in using the weak field equations that were similar to the renormalization problem to be described in Chapter 6.

> **While the theory of electro-weak interaction… was a great step forward, the research community at first found it difficult to accept. When they tried to use the theory for calculating in more detail the properties of the new W and Z particles (and many other physical quantities) it gave unreasonable results… Many researchers were pessimistic about the possibilities of going further with such a theory. — (Nobel press release, 1999)**

The solution to the problem was found in 1971 by two Dutch physicists, Martinus Veltman (b. 1931) and Gerardus ‘t Hooft (b. 1946). Of course the mathematics is way beyond our scope. Suffice it to say that they found a way to modify the equations so that meaningful calculations could be made. When their calculations were experimentally confirmed in 1995, Veltman and ‘t Hooft were awarded the 1999 Nobel Prize in physics:

> **The theory of the electro-weak force predicted the existence of the new W and Z particles right from the start. But it was only through ‘t Hooft’s and Veltman’s work that more precise prediction of physical quantities involving properties of W and Z could start. Large quantities of W and Z have recently been produced under controlled conditions at the LEP accelerator at CERN. Comparisons between measurements and calculations have all the time showed great agreement, thus supporting the theory’s predictions. – (Nobel press release, 1999)**

On this happy note we end the century-long saga that began in February 1896 when a French physicist couldn’t wait for a sunny day.

SUMMARY

The discovery of radioactivity eventually led to the discovery of a new quantum field — the *weak field,* tucked away in the tiniest nooks of the nucleus. It is this field that causes a neutron to change into a proton while emitting an electron and a neutrino — a process known as beta decay.

The weak field has a mass about 500 times greater that that of the strong field, so its range is about 500 times shorter. Like the strong field, the weak field has charge and comes in three versions, symbolized by W^+, W^- and Z°. Unlike the strong field, the weak field has spin 1, which puts it in the same family as the EM field; this relationship is called *electroweak unification*. However, the weak field is the only field that violates the law of parity – that is, its behavior is not left-right symmetric.

Because they have mass and charge, quanta of the strong and weak fields act even more like particles than do photons. In the next chapter we will see that matter fields with half-integral spin narrow the gap even further. The five force fields of nature are summarized in the following table. (Because the Higgs field is so new, I have not assigned a “color” to it.)

Field	**Spin**	**Quantum**	**Interacts with**	**“Color”**
Gravity	2	Graviton	everything	Blue
EM	1	Photon (γ)	charged fields	Green
Strong*	0	Pion (π^+, π^-, π°)	baryons	Purple
Weak	1	Weakon (W^+, W^-, Z°)	leptons and baryons	Brown
Higgs	0	Higgs boson	everything	?

*We will see in Chapter 10 that the strong field is made of more basic but “invisible” fields called quarks and gluons.

CHAPTER 6

MATTER – The "Red" and "Yellow" Fields

> **Just as there is an electromagnetic field, whose energy and momentum come in tiny bundles called photons, so there is also an electron field, whose energy and momentum and electric charge are found in the bundles we call electrons, and likewise for every species of elementary particle. The basic ingredients of nature are fields; particles are derivative phenomena. – S. Weinberg (W2001, p. 221)**

Now, finally, we come to matter itself. If you have been bored or bewildered by the arcane weak field and its mathematical subtleties, I urge you to pay attention as we come to the main course – the very quintessence of QFT: the *matter fields*. I use the plural because, as you will see, there are two matter fields: one for light particles (leptons) and the other for heavy particles (baryons).

Acceptance of matter as a field will not come easily. Think how difficult it was for physicists to accept force fields as having a real existence without requiring a mysterious ether. Well, giving up the concept of matter as "solid stuff" and accepting it as a pure field will be even harder. Even today, many physicists cling to the particle picture, as witness the common phrases "particle physics" and "elementary particles". Only a few, like Nobel laureates Frank Wilczek and Steven Weinberg, accept the pure QFT view that everything — even matter — is made of fields. But first I should explain what physicists mean by matter.

What's the matter? Matter is the stuff around us. It includes solids, liquids and gases. It is the stuff we can touch or feel (and also gases that we can't touch or feel). Matter is made of atoms that in turn are made of smaller entities called electrons, protons and neutrons. While this picture is now taught in elementary schools, there was a time when it was not easy to accept. (See Eddington's description of his writing desk in Chap. 1.)

Well now we must probe deeper and ask the question: What are these elementary particles made of? Are they the “hard, impenetrable, movable particles” that Newton spoke of? If so, what is inside them? Or are they point particles – things that have no spatial extent but nevertheless are somehow there?

The answer, according to QFT, is that particles are not solid balls or points; they are fields — properties of space, like the force fields. To be more specific, they are quanta of fields. In Chapter 3 we saw that Einstein’s concept of the photon as a particle defeated Planck’s field concept in the first round of the battle, even though Einstein later changed his mind:

> **In electrodynamics the continuous field appears side by side with the material particle as the representative of physical reality. This dualism, though disturbing to any systematic mind, has today not yet disappeared… The successful physical systems that have been set up since then represent rather a compromise between these two programs, and it is precisely this character of compromise that stamps them as temporary and logically incomplete… I incline to the belief that physicists will… be brought back to the attempt to realize that program which may suitably be called Maxwell’s: the description of physical reality by fields which satisfy… a set of partial differential equations. – A. Einstein (P1982, p. 289, 463)**

Particles vs. fields (review). Before proceeding, let us review the difference between particles and fields. A particle is a tiny solid object that travels through space like a bullet, or just sits there like a pea. A field is a condition or property of space that is found, at least to some extent, everywhere. To put it another way, a particle has boundaries, a field does not. To describe a particle, all one needs are its location and velocity, plus an equation stating how those two quantities change with time. To describe a field, one needs to know values of the field at every point of space, plus an equation showing how those values change with time. These equations are called *partial differential equations* and they determine how the field evolves and propagates through space.

We must remember, however, that quantum fields are not the same as classical fields; they are quantized. The matter field is not a single continuous field; it consists of chunks, or quanta, and each quantum has a life and death of its own. A quantum of field may spread over great distances, but it always maintains its separate identity. In QFT, what we call electrons, protons, etc., are really quanta of a field.

The first evidence for the field nature of matter came not from the nucleus, as it did for the strong and weak force fields, but from the electrons around the nucleus. These electrons are the “outside” of the atom, the face it presents to the world. It is the electron configuration that determines the chemical, optical, and physical properties of the atom – how it interacts with other atoms to form molecules, how it emits and absorbs light of different colors, etc. The job of the nucleus is to keep the electrons in place; the electrons do everything else. The story of how electrons came to be recognized as fields is a long and unfinished one. It began when Rutherford’s picture of orbiting electrons ran into problems.

PROBLEMS WITH ORBITS

An orbit, whether it be that of an electron around a nucleus or the moon around the earth, is the result of a balance between the tendency to keep moving in a straight line (inertia) and an attractive force that pulls the object inward (cf. Fig. 2-2). In the Rutherford atom the attractive force is electric, created by the positively charged nucleus. This picture soon ran into difficulties.

Stability. The first problem was the stability of the orbits. According to Maxwell's equations, an electron moving in an orbit radiates energy in the form of EM waves. As it loses energy it will move closer to the nucleus, just as an orbiting spacecraft fires retro-rockets when it returns to earth. We saw a gravitational example of this in Chapter 2, in which two stars spiraled inward because of energy lost by gravitational radiation. Thus the energy lost by EM radiation from an orbiting electron should cause it to spiral inward and eventually crash into the nucleus. But this doesn't happen.

Conclusion: Electrons can't be particles in orbit.

Discrete radiation. The second problem was the discreteness of EM radiation, demonstrated by Planck and Einstein. It doesn't take a genius (or does it?) to see that if light is emitted and absorbed in discrete amounts, then the atoms that do the emitting and absorbing must change their energy by discrete amounts. This would mean that electron orbits must exist as discrete energy states, and when a photon is emitted, the electron would jump from one orbit to another. But orbits do not change in jumps.

Conclusion: Electrons can't be particles in orbit.

Chemical properties. It has long been known that chemical properties are the same for all atoms of a given substance. For example, no oxygen atom behaves differently from other oxygen atoms — they are all the same chemically. But if the electrons were in classical orbits, variations would be inevitable.

Conclusion: Electrons can't be particles in orbit.

Spectroscopy. Finally, there was the problem with Fraunhofer lines, named after Joseph von Fraunhofer who chronicled them in 1814. They are the dark bands seen when sunlight is passed through a prism to separate it into its spectral components, or frequencies (Fig. 6-1).[1] These lines are caused by absorption of light, with each absorbing substance having its own specific frequencies of absorption. This would be impossible if electrons are particles in orbit.

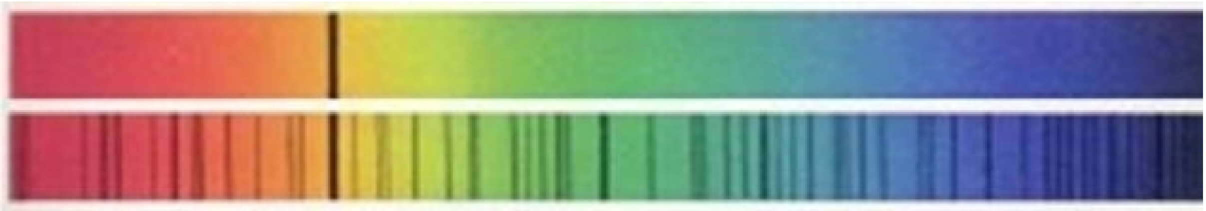
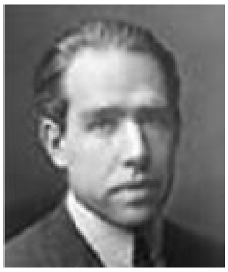


Fig. 6-1. (Top) White light from a lamp that has passed through sodium gas and is then spread out by a prism to display its component frequencies. The dark band is due to absorption by the sodium. (Bottom) Sunlight that has passed through a prism. The dark band at the same position as the sodium line above shows that sodium, among other elements, is present in the solar atmosphere.

Conclusion: Electrons can't be particles in orbit.

Yet the concept of particles in orbit was not so easily abandoned. The idea of solid particles was too ingrained, too intuitive. It was a Danish physicist who took the bull by the horns and announced that something new was needed.



Niels Bohr (1885-1962). After earning his doctorate in physics at Copenhagen University in 1911, Bohr went to England, where he worked at Cambridge University with J.J. Thomson, discoverer of the electron. Later Bohr joined Rutherford at Manchester University, shortly after Rutherford's discovery of the atomic nucleus. He soon came to the conclusion that Rutherford's picture of particles in orbit was untenable. In 1913 Bohr proposed that this view must be replaced by some new undefined electron states that satisfy the following two postulates:

> **1. [They] possess a peculiarly, mechanically unexplainable, stability.**
> **2. In contradiction to the classical EM theory, no radiation takes**

place from the atom in the stationary states themselves, [but] a process of transition between two stationary states can be accompanied by the emission of EM radiation. – N. Bohr (Nobel lecture, 1922)

Bohr did not say what the new states are, nor did he completely relinquish the idea of orbits. In fact, using Planck's relationship between energy of radiation and frequency, he showed that the frequencies of light emitted by hydrogen can be derived from the assumption of stable orbits — even though such orbits are not possible! His calculation was aided by the fact that Johann Balmer in 1885 had discovered an empirical equation for the Fraunhofer frequencies of hydrogen. Balmer's formula contained an arbitrary constant, and what Bohr did was to show that Balmer's empirical constant can be derived from a hybrid picture, showing that there was a basis for a true theory.

Quantum leaps. The discontinuous transitions that Bohr postulated soon came to be known as "quantum leaps". Some of you may wonder, if the quantum of energy is so small, why do we use the phrase "quantum leap" in everyday speech to indicate a big change? The answer lies not in the magnitude, but in the idea of a discontinuous jump, as opposed to a gradual change. Yes, the quantum of energy is exceedingly small by human standards, but the significance in physics of the word quantum is not its smallness but its discreteness. In the atomic world things occur in jumps or steps. It's like the difference between a ramp and a stairway. On a ramp you can creep your way up by whatever small amounts you choose, but on a stairway you can only go one step at a time. (And have you older readers noticed how these steps keep getting higher?) Going up a step is like making a quantum leap.

Following Bohr's success, there was a rush to apply his seed of a theory to atoms other than hydrogen, but soon another difficulty arose — that was solved by Wolfgang Pauli.

Wolfgang Pauli (1900-1958). Pauli was born in Vienna in the same year that

Planck introduced quantization. He was five when Einstein published his theory of relativity. At the age of 19, while a student at the University of Munich, he was asked to write an encyclopedia article on Einstein's theory. The article was so brilliant that when Einstein read it he commented that perhaps Pauli knew more about relativity than he himself did. In 1925, five years before postulating the neutrino (Chapter 5), Pauli showed that the various atomic spectra would make sense if each electron is in a different state from the other electrons. Pauli called this assumption the *Exclusion Principle*.

The Exclusion Principle. This principle states that each electron in an atom can be described by four *quantum numbers* and each state can be occupied by only one electron. Once a given state is filled, other electrons are excluded from that state. Pauli realized that one of the numbers represents energy, which is related to the distance from the nucleus, while two other numbers represent angular momentum, which has to do with the shape and orientation of the orbit. However, Pauli could find no physical significance for the fourth quantum number, although it was needed to fit the data. The significance was found that same year by two Dutch physics students.

Electron spin. George Uhlenbeck and Samuel Goudsmit were studying details of spectral lines known as the *anomalous Zeeman effect*. This eventually led them to the realization that Pauli's fourth quantum number must relate to electron spin. Here is the story told in Goudsmit's delightful and self-effacing words — a story that conveys the groping and uncertainty that exist in the struggle to understand nature, as contrasted with the logic and certainty that are imposed after the battle is won.

> **The Pauli [exclusion] principle was published early in 1925… if I had been a good physicist, then I would have noticed already in May 1925 that this implied that the electron possessed spin. But I was not a good physicist and thus I did not realize this… Then Uhlenbeck appears on the scene… he asked all those questions I had never asked… When the day came that I had to tell Uhlenbeck about the Pauli principle — of course using my own quantum numbers — then he said to me: "But don't you see what this implies? It means that there is a fourth degree of freedom for the electron. It means that the electron has a spin, that it rotates"… I asked him: "What is a degree of freedom?" In any case, when he made his remark, it was luck that I knew all these things about the spectra, and I said: "That fits precisely in our hydrogen scheme**

which we wrote about four weeks ago. If one now allows the electron to be magnetic with the appropriate magnetic moment, then one can understand all those complicated Zeeman effects. — S. Goudsmit (www.lorentz.leidenuniv.nl/history/)

And so was introduced the idea that the electron spins on its axis (still thinking of the electron as a particle) and that this spin has a value of ½ Planck units, as contrasted with the photon's spin of 1.

Confirmation. The next step was to see if this idea is consistent with experiment. During a course at Harvard University, Prof. Wendell Furry gave the following whimsical account of how this happened (as best I recall it):

After Uhlenbeck and Goudsmit had the idea that electron spin might explain the anomalous spectroscopic results, there remained the crucial task of determining if the effect is in the right direction. This involved the kind of calculation that all physics students have to suffer through, in which polarity, direction of spin, direction of magnetic field, the "right-hand rule", etc., get all confused and make the head spin (no pun intended). In other words, there are many opportunities to go wrong, and many do. Well, the story goes, each man did the calculation and when they compared notes they found they had opposite results. One of them had obviously made a mistake, so they went back to check their calculations. As it happened, each found an error, so they were still in disagreement. At this point they went to their mentor, Paul Ehrenfest, who happened to have a distinguished visitor named Albert Einstein. It was decided that the four of them would do the calculation independently (remember, we are talking about elementary physics here). When they got together the result was now 2-2. They finally broke the deadlock by counting Einstein's vote twice. — W. Furry (reconstructed)

I'm sure this story was a joke, making fun of the difficulty physicists have in keeping track of the proper sign. (It is said that the difference between a good and a bad physicist is that the good one makes an even number of errors, so the final sign comes out right.) It also pokes fun at the "papal" authority of Einstein. However, Uhlenbeck did say that Einstein visited Leiden in 1925 and "gave us the essential hint" to complete the calculation (Physics Today 29, Dec. 1976, p. 6).

Pauli's Exclusion Principle, important as it was, did not solve the problem of electron orbits. If they are not particles in orbit, what are they? The first step toward an answer was made by a French aristocrat who spent World War I in the Eiffel Tower.

Louis de Broglie (1892-1987). Prince Louis de Broglie was the son of a noble French family. His great-great-grandfather had the distinction of being beheaded in the French revolution. After majoring in literature and history, de Broglie changed to physics, perhaps influenced by his older brother Maurice, who was among the 24 scientists invited to the 1911 Solvay Conference. Upon returning from that meeting, Maurice told his brother about the excitement generated by the dilemma of the photon: a wave that behaves like a particle. Two years later, de Broglie heard of Bohr's dilemma: that the electron, long thought of as a particle, doesn't behave like a particle. He wondered if there might be a connection between the two dilemmas:

> **The necessity of assuming for light two contradictory theories – that of waves and that of corpuscles – and the inability to understand why, among the infinity of motions which an electron ought to be able to have in the atom according to classical concepts, only certain ones were possible: such were the enigmas confronting physicists at the time I resumed my study of theoretical physics. – L. de Broglie (Nobel lecture, 1929)**

But before he could resume his study, World War I had broken out and Prince Louis was conscripted into the French army (*noblesse oblige*) where he served four years stationed in the Eiffel Tower as a radio operator.

> **When in 1920 I resumed my studies of theoretical physics which had long been interrupted by circumstances beyond my control, I was far from the idea that my studies would bring me several years later to receive such a high and envied prize as that awarded by the Swedish Academy of Sciences… I was attracted to theoretical physics by the mystery enshrouding the structure of matter and the structure of radiations, a mystery which deepened as the strange quantum concept introduced by Planck in 1900 in his**

research on black-body radiation continued to encroach on the whole domain of physics.– L. de Broglie (Nobel lecture, 1929)

Matter waves. In 1920 de Broglie came up with the idea that would bring him the Nobel Prize. If electrons, like photons, exhibit both wavelike and particle-like behavior, he thought, it would bring a satisfying unity to nature. Of course if he had stopped there, he might as well have stayed in the Eiffel Tower; philosophical speculation won't buy you a footnote, let alone a Nobel Prize. De Broglie went on to show that a wave picture of the electron would indeed explain Bohr's mysterious discrete energy states. Well, it if the electron is a wave, there must be something to do the waving. So at this time, ladies and gentlemen (fanfare please), I would like to introduce…

THE MATTER FIELDS

In fact, I would like to introduce two matter fields: one for the electron and other leptons, and another for the proton and other baryons. The two fields differ primarily in their masses and in their interactions with other fields. Like the force fields, matter fields are properties of space that are present everywhere, and their behavior is governed by partial differential equations that describe how the field strength at each point is affected by the field strength at adjacent points. The electron is a quantum of the lepton field just as the photon is a quantum of the EM field. Gone are sharp edges, round balls and point particles. In QFT, the electron, proton, and other “particles” are chunks of field that spread out in space but nevertheless act as units. Whenever I use the word *particle*, I really mean *field quantum*.

> **The reason nature has a particle-like aspect is that it is made of fields and these fields are quantized. Although it’s legitimate to think of electrons, protons, and so on as particles, it’s important to remember that they are not particles in the simple Newtonian sense. An electron is nothing like, say, a tiny pea. An electron is simply an energy increment of a spread-out matter field. – A. Hobson (H2007, p. 312)**

Color them yellow and red. In the color scheme that I’m using to help you visualize the fields of nature, I choose yellow for lepton fields (electrons, neutrinos, etc.) and red for baryon fields (protons, neutrons, etc.). As always, it is important to remember that these colors are not real; they are only a way to imagine properties of space that do not present themselves directly to our senses.

What about Rutherford? This brings up the question, if the electron is a field that fills the space around the nucleus, how could Rutherford conclude that gold atoms are mostly empty space? The answer is that the electron field is not a solid impenetrable object; it is a condition of space, and different fields can be present in the same space. In the same “yellow” space that represents the electron field around a nucleus, there is also a green EM field and a very weak blue gravitational field. There are even slight tinges of the short-range purple and brown fields that are present inside the nucleus. So it was no problem for Rutherford’s alpha ray to barrel right through these diffuse fields until it is deflected by a highly-concentrated nucleus, where the electrical repulsive force is much stronger.

The atom revisited. Fig. 6-2 is a simplified picture of the electron field in a hydrogen atom. Instead of particles in orbit (Fig. 4-1), the electron is seen as a "yellowness" of the space surrounding the nucleus, dropping off in intensity at greater distances, but never becoming zero. (The nucleus, here shown as a red dot, is also made of fields whose intensities fade away very rapidly.)

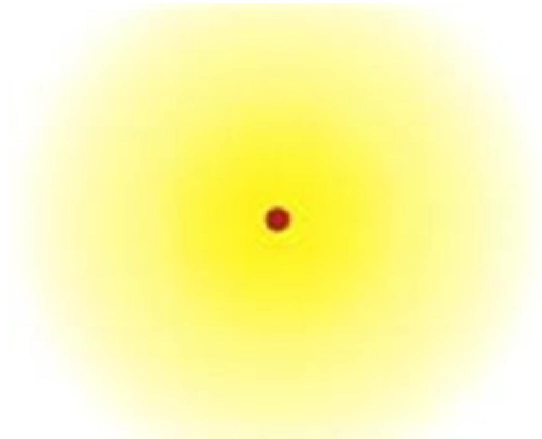


Fig. 6-2. A simplified picture of a hydrogen atom, showing the nucleus (red) surrounded by an electron field (yellow).

Bound but not attached. While the yellow electron field in Fig. 6-2 resembles the green EM field around the nucleus (Fig. 3-1a), there is an important difference. The electron field is not attached the way the EM self-field of an electron is attached. The electron is a true field quantum with an independent existence. It may be held in place temporarily (or even indefinitely) by the electrical attraction of a nucleus, but given enough energy it is perfectly capable of going off on its own. It is a quantum that is temporarily bound, like a bird in a cage, but if the door is left open it can go flying off.

Oscillations. The static field picture shown in Fig. 6-2 does not in itself explain the discrete energy states. We must now add oscillations — an alternating increase and decrease of field intensity. However, these fluctuations are not synchronous with all points pulsating in unison. The field equations do not permit that, any more than a water surface can simultaneously move up and down everywhere at the same time. Instead, the oscillations form a pattern of waves, as shown in Fig. 6-3 (again, highly simplified). As the intensities oscillate, these waves travel around the nucleus.

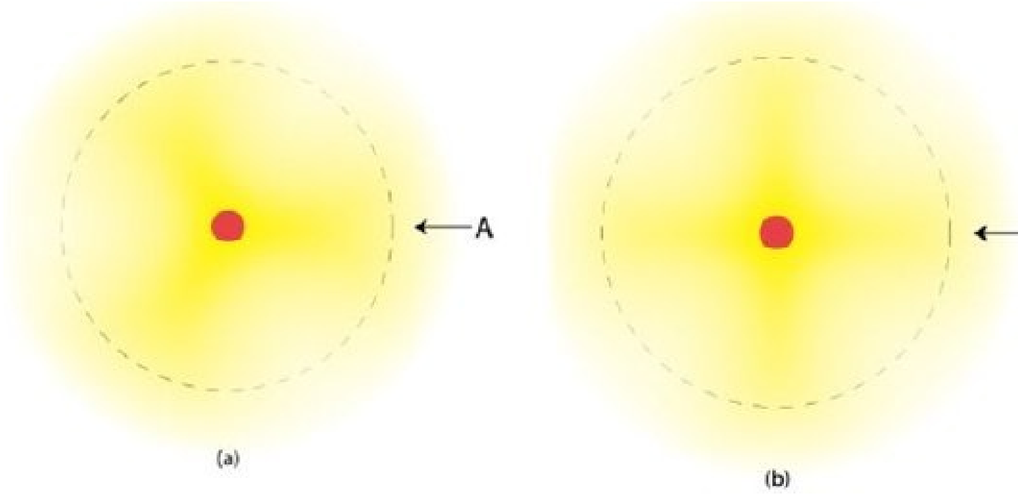


Fig. 6-3. Two simplified patterns of oscillation of the electron field in a hydrogen atom. (a) Three intensity peaks surround the nucleus. (b) With a higher frequency of oscillation, there are four peaks. These patterns change with time as the fields oscillate. Each diagram is a "snapshot" taken at a single instant.

Discreteness explained. By studying the patterns in Fig. 6-3, you can see why there must be an integral number of waves. As we travel around the dashed circles, the field intensity oscillates but it must return to its original value at the starting point. This is illustrated in Fig. 6-4, which show plots of field intensity along the dashed circles in Fig. 6-3. In (a) and (b), the intensity returns to the same value at point A, after passing through two or three peaks, but if we choose an intermediate frequency as in (c), the pattern is impossible because the intensity at the starting point A cannot have two different values at the same time.

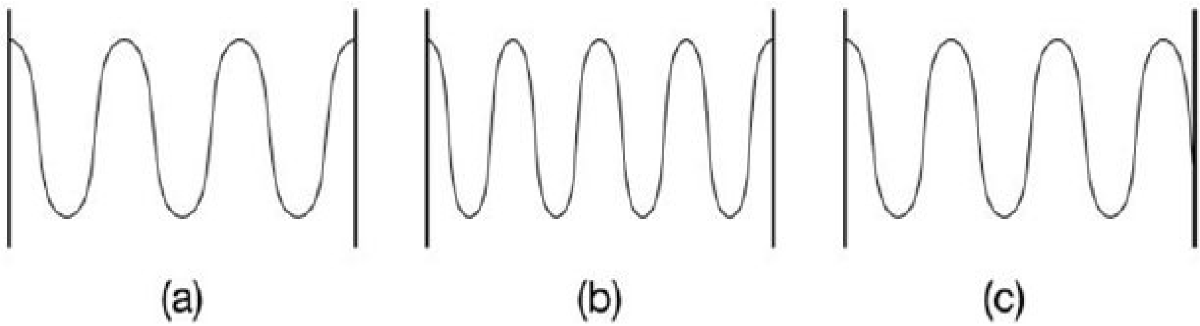


Fig. 6-4. (a) and (b) are plots of field intensity along the dashed circles in Fig. 6-3 (a) and (b). (c) is the same plot for a frequency of oscillation intermediate between (a) and (b)

This was de Broglie's great insight: if the electron in an atom is made of waves then the number of waves must be an integer, and the corresponding frequencies must be discrete. And since the frequency of oscillation is related to the energy of the electron field, the energy states must be discrete.

> **Determination of the stable motion of electrons in the atom introduces integers… That suggested the idea to me that electrons themselves could not be represented as simple corpuscles either, but that a periodicity had also to be assigned to them too… The conditions of quantum stability thus emerge as analogous to resonance phenomena and the appearance of integers becomes as natural here as in the theory of vibrating cords and plates. – L. de Broglie (Nobel lecture, 1929)**

A musical analogy. The resonance phenomena that de Broglie referred to are commonplace in music. When sound waves are produced in a bounded space, e.g., a violin string or a trombone, only certain frequencies are able to maintain themselves, depending on the length or size of the space. (There are also two-dimensional examples, like tympani.) To change the frequency or pitch, one must either change the dimensions of the space or the properties of the vibrating medium. The former happens when the violinist moves his finger along the string to alter the effective length or when the trombone player moves the slide. The latter occurs, for example, when a violin and tympani are tuned by changing the tension, or when the trombone goes sharp because the air temperature rises.

For the electron field around a nucleus the situation is more complex because it is three-dimensional and there is no physical boundary, but the principle is the same: Only certain frequencies can sustain themselves in the available space. This field picture of an electron fulfills Bohr's postulates and provides a simple answer to the problems with orbits described above.

A prediction. If de Broglie had stopped there it would have been enough, but he went on to apply his wave idea to free electrons. If bound electrons are associated with waves, he reasoned, then free electrons must also be associated with waves and these waves would produce interference effects, as do photons (see Fig. 3-5). Using Planck's relationship between frequency and energy, de Broglie calculated the wavelength (distance between wave peaks) of a moving electron and found it to be comparable to that of x-rays. Since x-ray interference had been demonstrated by using the interatomic spaces in a crystal in lieu of slits (because of the shorter wavelength of x-rays), de Broglie suggested that the same interference patterns could be seen using a beam of electrons:

> **For x-rays the phenomenon of diffraction by crystals was a natural consequence of the idea that x-rays are waves analogous to light and differ from it only by having a smaller wavelength. For electrons nothing similar could be foreseen as long as the electron was regarded as a simple small corpuscle. However, if the electron is assumed to be associated with a wave… then a [similar] phenomenon ought to be expected… If the phenomenon actually exists, it should thus provide decisive experimental proof in favor of the existence of a wave associated with the electron… and so the fundamental idea of wave mechanics will rest on firm experimental foundations. – L. de Broglie (Nobel lecture, 1929)**

De Broglie presented these ideas in his Ph.D. thesis of 1924, but they were so startling, so unexpected, that his thesis advisor, Paul Langevin, didn't want to accept the thesis. He did, however, send a copy to Albert Einstein, who praised the work effusively, saying that it "lifted the corner of the great veil". That was enough for Langevin, and the thesis was quickly accepted.

Confirmation. De Broglie's suggestion that electrons can produce diffraction effects was tested by Clinton Davisson at Bell Laboratories in New York and by George Thomson (son of J. J. Thomson, discoverer of the electron) at the University of Aberdeen. The results were as de Broglie had predicted. Beams of electrons reflected from crystalline structures showed the same bright-dark pattern that was seen in similar optical and x-ray experiments. What's more, the wavelength calculated from the experiment agreed with de Broglie's prediction! Following this, de Broglie received the 1929 Nobel prize and Davisson and Thomson shared the 1937 award.

Further validation was obtained when the famous two-slit experiment with photons (see Fig. 3-5) was replicated using electrons — an experiment that Richard Feynman said couldn't be done.

> **Now we imagine a similar experiment with electrons… We should say right away that you should not try to set up this experiment. This experiment has never been done in just this way. The trouble is that the apparatus would have to be made on an impossibly small scale to show the effects we are interested in. – R. Feynman ("The Feynman lectures on Physics", 1963)**

The "impossible" experiment was performed in 1961 by Claus Jőnnson of the

University of Tübingen in Germany, and it was repeated in 1974 by Italian physicists Gian Franco Missiroli, Giulio Pozzi and Pier Giorgi Merli using single electrons. This latter experiment was voted “the most beautiful experiment” by readers of *Physics World*, but no Nobel prizes were awarded for either one. The results of these experiments are shown in Fig. 6-5.

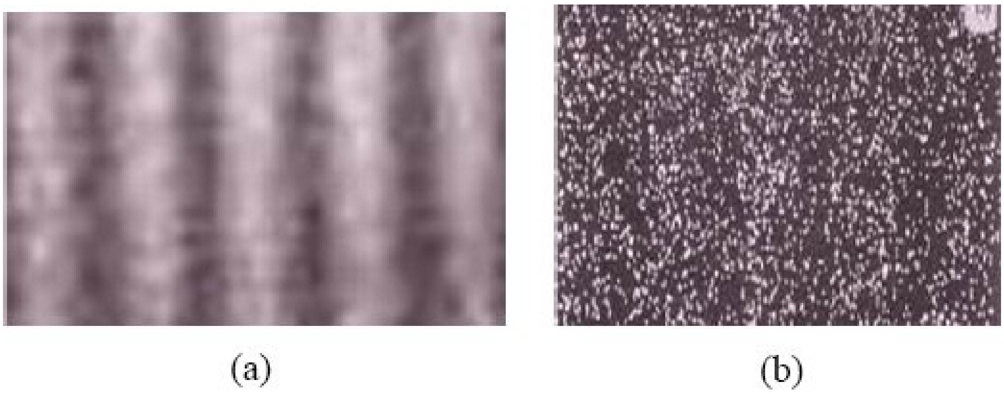


Fig. 6-5. The two-slit experiment with electrons. (a) The pattern after many electrons are detected. (b) The pattern after a small number of electron impacts; each dot was created by a single electron. (H2007, p. 309)

Particles persist. Given all this success, you might think that the particle picture was dead and the electron was now established as a field. You would be wrong. Despite the wave behavior shown in Fig. 6-5, the particle concept held on. The idea that an electron field could collapse into a single atom was too much for people to accept (see “quantum collapse” below). De Broglie himself was not willing to abandon the idea of the electron as a particle. This insistence on particles led to a theory known as…

Quantum Mechanics. As its name indicates, Quantum Mechanics is a theory of particles. Their behavior is described by a *wave-function* that gives the *probability* that a particle is at a given location and whose evolution is described by Shrödinger’s equation. Ironically, the man who derived this equation believed (or wanted to believe) that it describes fields, not particles.

THE EQUATION

Field equations describe how the field intensity at one point is related to the intensity at another point and thus how the field propagates through space. Gravity has its Einstein equation (Chapter 2), EM its Maxwell equations (Chapter 3), the strong field has the Yukawa equation (Chapter 4), and the weak field has the equation with many authors (Chapter 5). If the electron and other particles of matter are really fields, there must be partial differential equations that describe their behavior. At least so thought Erwin Schrödinger.

Erwin Schrödinger (1887-1961). While the young de Broglie was serving France in the Eiffel Tower, an Austrian physicist named Erwin Schrödinger was serving on the other side as an artillery officer on the Italian front. These two war-time enemies were destined to become close friends and allies in creating the new physics known as quantum mechanics. After the war Schrödinger wound up at Zürich, where Einstein had developed his theory of relativity. In 1927 he took over the Max Planck chair at the University of Berlin, where he remained until Hitler came to power.[2] Schrödinger learned of de Broglie's work by reading a footnote in a paper by Einstein and even before the experimental confirmation, he realized its significance:

> **I have been intensely concerned these days with Louis de Broglie's ingenious theory. It is extraordinarily exciting, but still has some very grave difficulties. – E. Schrödinger (letter to Einstein, Nov. 3, 1925)**

One day while Schrödinger was giving a seminar a student asked him why there was no equation to show how de Broglie's waves change with time, as there was for EM waves. Schrödinger got to work and in two weeks found his now famous equation, which he published in 1926. With a field equation in hand, Schrödinger did not see any need for an accompanying particle, as did de Broglie. To Schrödinger the electron was a field and only a field.

Another field with mass. Like the equations for the strong and weak fields, Schrödinger's equation for the electron contains a mass term that slows down the propagation of changes. Without this term, matter fields would travel at the speed of light, but with the mass term they propagate more slowly — in fact, like a particle of mass m. Another effect of mass is to increase the frequency of oscillation, which you may recall is related to energy. In fact, this relation between mass and energy leads directly to the most famous equation in physics, as we will see in Chapter 10.

And spin. As we have noted, the electron has a spin of ½. We have learned that in QFT spin (or helicity) is related to the number of internal components, and as it turns out, a spin ½ field must contain four components. Schrödinger's equation, however, describes only a single number called the wave function. Just as with EM and gravity, an equation was needed that contains multiple components.

Paul A.M. Dirac (1902-1984). The problem that first concerned Dirac was not so much the spin, but the fact that Schrödinger's equation did not conform with Einstein's theory of relativity (App. A). This theory, published in 1905, was based on the principle that the equations of physics are the same for all observers, regardless of their state of motion so long as it is uniform. As it happens, Oskar Klein and Walter Gordon had already suggested a relativistic equation with one component (see footnote 6 in Chapter 5). Dirac's breakthrough was to expand this number to four in order to accommodate the spin ½ of the electron. His result was published in 1928, and the matter fields are sometimes referred to as Dirac fields.

At first the Dirac equation was applied only to the electron, but when it was learned that the proton and neutron also have spin ½,[3] the same equation was used to describe their behavior, but with different masses and different interaction terms.

Fields vs. Particles (Round 2). Dirac and Schrödinger shared the 1933 Nobel Prize in physics, but their views on the equations they had fathered were quite different. Dirac believed that electrons are particles and his equation was a mathematical tool for calculating the probability of finding a particle at a given point:

> **Dirac thought that… "the role of the field is to provide a means for making observations of a system of particles" and therefore "we cannot suppose the field to be a dynamical system on the same footing as the particles and thus be something to be observed in the same way as the particles". – S. Tomonaga (Nobel lecture, 1965)**

Schrödinger, on the other hand, wanted to believe that the electron is a wave spread throughout all of space, but he had to face the particle behavior as seen, for example, in a cloud chamber (Fig. 6-6).

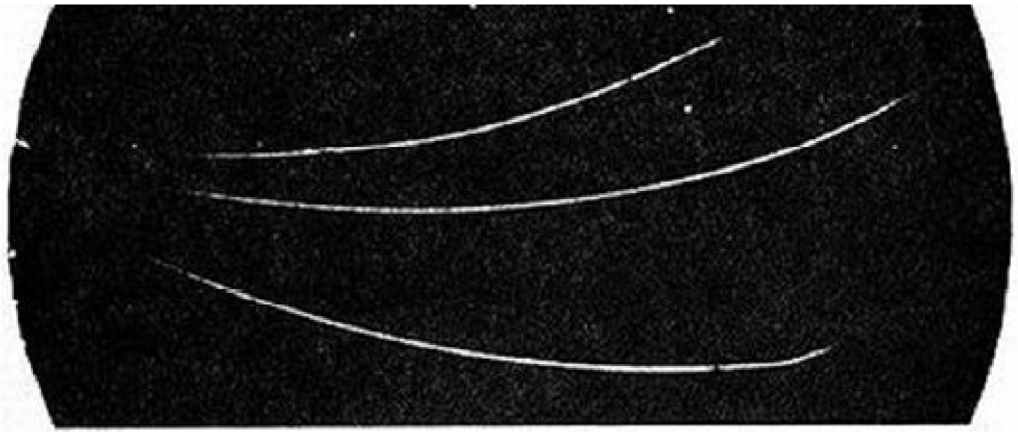


Fig. 6-6. Tracks left by alpha particles as they traverse a cloud chamber. The curvature is due to an applied magnetic field. (www.practicalphysics.org)

> **We assert that the atom in reality is merely the… phenomenon of an electron wave captured, as it were, by the nucleus of the atom… From the point of view of wave mechanics, the [particle picture] would be merely fictitious. I have, however, already mentioned that we have yet really observed such particle paths… We find it confoundedly difficult to interpret the traces we see as nothing more than narrow bundles of equally possible paths. – E. Schrödinger (Nobel lecture, 1933)**

Quantum collapse. The answer to Schrödinger's dilemma is quantum collapse (see Chapter 3). If a field quantum transfers its energy to an atom it must do so as a unit and disappear from all space, no matter how spread out it is. Even if it transfers only part of its energy, it must collapse every time energy is transferred. As Art Hobson wrote, referring to the cloud chamber tracks shown above:

The tracks are made by successive individual interactions between a matter field and gas or water molecules. The matter quantum collapses… each time it interacts with a molecule, while spreading out as a matter field between impacts. – A. Hobson (H2007, p. 313)

Particles win… again. The outcome of this second battle was another victory for particles and another defeat for fields. For one thing, Dirac's textbook (D1958) had become the Bible. For another, Werner Heisenberg had developed an alternate but mathematically equivalent version of quantum mechanics that didn't involve fields. Finally, and perhaps most importantly, quantum collapse was not accepted, or even suggested, as an explanation for the particle-like behavior of fields. Most physicists, ignoring Schrödinger and de Broglie, chose to believe in QM — a theory in which reality is made of particles, described by a wave function that gives the probability of finding the particle at a given point.

RENORMALIZATION

The Dirac equation soon ran into calculational problems, related to the way a charged field quantum - say an electron - interacts with its self-field (see Fig. 6-7). According to the equations of QFT, this interaction affects properties of the electron, such as its effective mass and charge. When these properties are calculated, the results turn out to be infinite. The unavoidable conclusion is that the QFT equations do not describe these self-interactions correctly.

> **The first phenomenon which attracted our attention as a manifestation of field reactions was the electromagnetic mass of the electron. The electron, having a charge, produces an electromagnetic field around itself. In turn, this field, the so-called self-field of the electron, interacts with the electron. We call this interaction the field reaction. Because of the field reaction the apparent mass of the electron differs from the original mass. The excess mass due to this field reaction is called the electromagnetic mass of the electron and the experimentally observed mass is the sum of the original mass and this electromagnetic mass … But what will happen if the influence of field reaction is taken into account? This theoretical problem was examined… [and] the influence of the field reaction becomes infinite… This does not, of course, agree with experiment. This discouraging state of affairs generated in many people a strong distrust of quantum field theory. – S. Tomonaga (Nobel lecture, 1965)**

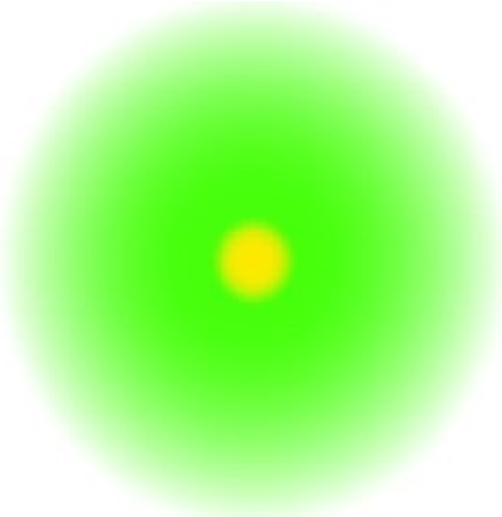


Fig. 6-7. Self-field of an electron. The "green" field surrounding the electron is permanently attached; it is called the self-field. According to the equations of QFT, this self-field acts on the electron and affects its dynamics.

The solution, called *renormalization*, proved wildly successful. The finest details of electromagnetic interactions could now be calculated to an unheard-of accuracy. The solution was found by three people, working independently, who shared the 1965 Nobel Prize.

Sin-Itiro Tomonaga (1906-1979). Tomonaga was the son of a philosophy professor at Kyoto University. He spent two years in Germany working with Werner Heisenberg and when he returned to Japan in 1939 this work became the basis for his doctoral thesis. During WWII Tomonaga worked on radar systems (as did his US counterpart Julian Schwinger). After the war, Tomonaga turned to the problem that was stumping physicists in the 1940's. He found a simple solution: Just replace the calculated values of mass and charge, infinite though they may be, with the experimental values.

> **Since those parts of the modified mass and charge due to field reactions [become infinite], it is impossible to calculate them by the**

theory. However, the mass and charge observed in experiments are not the original mass and charge but the mass and charge as modified by field reactions, and they are finite. On the other hand, the mass and charge appearing in the theory are… the values modified by field reactions. Since this is so, and particularly since the theory is unable to calculate the modified mass and charge, we may adopt the procedure of substituting experimental values for them phenomenologically… This procedure is called the renormalization of mass and charge… After long, laborious calculations, less skilful than Schwinger's, we obtained a result… which was in agreement with Americans'. – S. Tomonaga (Nobel lecture, 1965)

Julian Schwinger (1918-1994). Schwinger started early on his mission to develop quantum field theory. In 1934, at the incredible age of 16, he wrote an unpublished paper that attempted to extend to all fields what Dirac had done for the electron field. At 18, he received his Bachelor's degree from Columbia and the next year he published seven (!) articles on the properties of neutrons, including the first determination of neutron spin. These papers became his Ph.D. thesis, but he didn't receive the degree until 1939 because of Columbia's residence requirement. Schwinger spent the war years working on radar systems at the MIT Radiation Laboratory. After the war, he returned to the "mountain" of QFT that he started to climb at 16 — hence the title of his biography: "Climbing the Mountain" (M2000). This time his trip led to the renormalization solution. Schwinger's presentation of these results at the Annual Meeting of the American Physical Society in January 1948 was well-received:

The great event came on Saturday morning, and was an hour's talk by Schwinger, in which he gave a masterly survey of the new theory which he has had the greatest share in constructing and at the end made a dramatic announcement of a still newer and more powerful theory, which is still in embryo. This talk was so brilliant that he was asked to repeat it in the afternoon session, various

unfortunate lesser lights being displaced in his favour. There were tremendous cheers when he announced that the crucial experiment had supported his theory. — F. Dyson (M2000, p. 255)

By the time of the Pocono conference in March, Schwinger had further refined his method. Here is the report by C.N. Yang, illustrating again Schwinger's inability to communicate his ideas to the physics community:

Fermi took voluminous notes because he was aware that it was a historical event to listen to what Schwinger had to say. After they came back to Chicago… Fermi gathered Teller and Wentzel and four graduate students… into his office, and we spent weeks trying to digest what Fermi had written down… After about six weeks of meeting several times a week in Fermi's office for something like two hours each session we were all very tired, and none of us felt that we had understood what Schwinger had done. We only knew that Schwinger had done something brilliant. – C.N. Yang (N1996, p. 176)

Richard Feynman (1918-1988). Today Richard Feynman is probably the best-known of 20th-century American scientists, enshrined on a postage stamp. Among the reasons for his fame are several delightfully-written popular books (F1985, F1988) and the public role he played in solving the 1986 space shuttle disaster. Feynman spent the war years at Los Alamos working on the atomic bomb and figuring out how to open other people's safes. (This playful and harmless habit later caused him some grief when it was entered into his FBI file.) When it came to the fields-vs.-particles debate, Feynman was strictly a particle man:

We know that light is made of particles because we can take a very sensitive instrument that makes clicks when light shines on it, and if the light gets dimmer, the clicks remain just as loud — there are just fewer of them. Thus light is something like raindrops — each

> **little lump of light is called a photon… You might wonder how it is possible to detect a single photon. One instrument that can do this is called a photomultiplier… You might say that it's just the photomultiplier that detects light as particles, but no, every instrument that has been designed to be sensitive enough to detect weak light has always ended up discovering the same thing: light is made of particles. — R. Feynman (F1985a, p. 14-15)**

It did not occur to Feynman that these discrete clicks could also be produced by light quanta collapsing into atoms of the photomultiplier — or if it did, he dismissed the idea. By overlooking or ignoring this possibility, Feynman was forced to accept a theory that he himself admitted doesn't make sense:

> **It is not a question of whether a theory is philosophically delightful, or easy to understand, or perfectly reasonable from the point of view of common sense. The theory of quantum electrodynamics describes Nature as absurd from the point of view of common sense. And it agrees fully with experiment. So I hope you can accept Nature as She is — absurd. - R. Feynman (F1985a, p. 10)**

I consider it a tragedy that Feynman, perhaps the most influential physicist of his time, could not accept QFT, which not only is "philosophically delightful" but also "agrees fully with experiment"..

> **Feynman… hoped that by formulating his theory directly in terms of paths of particles… he would avoid the field concept and construct something essentially new. For a while he thought he had. Why did he want to get rid of fields? 'I had a slogan', he said… 'The vacuum doesn't weigh anything (dramatic pause) *because there's nothing there.*'" — F. Wilczek (W2008, p. 84)**

It is not surprising, therefore, that Feynman's initial approach to the self-interaction problem was based on particles, not fields:

> **Well, it seemed to me quite evident that the idea that a particle acts on itself, that the electrical force acts on the same particle that generates it, is not a necessary one — it is a sort of a silly one, as a matter of fact. And, so I suggested to myself, that electrons cannot act on themselves, they can only act on other electrons. That means there is no field at all. You see, if all charges contribute to making a single common field, and if that common field acts back on all the**

charges, then each charge must act back on itself. Well, that was where the mistake was, there was no field. It was just that when you shook one charge, another would shake later. There was a direct interaction between charges, albeit with a delay…[4] Shake this one, that one shakes later. The sun atom shakes; my eye electron shakes eight minutes later, because of a direct interaction across [the intervening space]. – R. Feynman (Nobel lecture, 1965)

When Feynman presented his method at the Pocono conference (see above), his speech was scheduled after Schwinger's.

He was uneasy. It seemed to him that Schwinger's talk had not gone well (but he was wrong – everyone, and crucially Oppenheimer had been impressed.)… He did have a mathematical formalism, as private though not as intricate as Schwinger's, and he could show how to derive his rules and methods from this formalism, but he could not justify the mathematics itself. He had reached it by trial and error. He knew it was correct because he had tried it now on so many problems, including all of Schwinger's, and it worked, but he could not prove that it worked. – J. Gleick (G1992, p. 257)

The conflict was resolved in 1949 by Freeman Dyson, who showed that Feynman's particle diagrams gave the same results as the field theory approach of Schwinger and Tomonaga.

Fields vs. Particles (Round 3). For the third time in physics history there had been a momentous battle between particles and fields, and for the third time fields lost and particles won. In 1905, Einstein's view of light as particles (later recanted) replaced Planck's field quantum picture (Chapter 3). In 1933, Dirac's particle theory of QM won over Schrödinger's field picture. And then in 1948, Feynman's approach to renormalization based on particles won over Schwinger's fields, in large part because his diagrams proved easier to work with than Schwinger's field equations. Even though Feynman eventually changed his mind (see "Feynman converts" below), two generations of physicists have been brought up on Feynman diagrams and believe that nature is made of particles.

One cannot fault Feynman for finding a method that works and is easy to use, nor can one fault him for being on the "particle" side, as was Dirac and, for a time, Einstein. However, I do fault him for his dismissal of QFT, an elegant and consistent theory that explains so many things, as a "shell game"

(F1985a, p. 6). I even heard him make sneering references to Schwinger's mathematical approach during a lecture before the American Physical Society. While "bringing computation to the masses", as Schwinger said in his eulogy for Feynman, Feynman gave them an admittedly absurd theory without any theoretical basis while ignoring a well-founded theory that provides a consistent, paradox-free picture of nature.

Feynman converts! According to Frank Wilczek, Feynman eventually lost confidence in his particles-only view of nature:

> **Feynman told me that when he realized that his theory of photons and electrons is mathematically equivalent to the usual theory, it crushed his deepest hopes… He gave up when, as he worked out the mathematics of his version of quantum electrodynamics, he found the fields, introduced for convenience, taking on a life of their own. He told me he lost confidence in his program of emptying space – F. Wilczek (W2008, p. 84, 89)**

Unfortunately, Feynman's "conversion" is not generally known. Most physicists today routinely use Feynman diagrams while promulgating and perpetuating the particle picture, puzzling and paradoxical as that picture may be.

Schwinger's third trip up the mountain. While renormalization was accepted by most physicists as being both legitimate and necessary, not everyone was happy with it:

> **Despite the experimental success of the theory… the fact that the infinities occur at all continues to produce grumbling… Dirac in particular always referred to renormalization as sweeping the infinities under the rug. I disagreed with Dirac and argued the point with him at conferences at Coral Gables and Lake Constance. — S. Weinberg (W1992, p. 115)**

Schwinger also was not happy with the situation, but for a different reason. He felt that the theory should include force fields and matter fields on an equal basis. This led to his third and final trip "up the mountain."

> **The pressure to account for those [experimental] results had produced a certain theoretical structure that was perfectly adequate for the original task, but demanded simplification and generalization… I needed time to go back to the beginnings of**

> **things… My retreat began at Brookhaven National Laboratory in the summer of 1949. It is only human that my first action was one of reaction. Like the silicon chip of more recent years, the Feynman diagram was bringing computation to the masses… But eventually one has to put it all together again, and then the piecemeal approach loses some of its attraction… Quantum field theory must deal with [force] fields and [matter] fields on a fully equivalent footing… Here was my challenge. – J. Schwinger (B1983, p. 343-345) [5]**

The result was six articles published in 1951-54, titled "The Theory of Quantized Fields". Schwinger believed that this achievement was far more important than the renormalization work for which he had received the Nobel Prize. Unfortunately, this work is rarely cited.

> **And so a tragedy ensued. [Schwinger's] need to do things his way made him develop his own language, his own approaches and techniques… As he became more isolated, fewer people understood and spoke the newer languages he created — for example, sourcery — contributing to his further isolation… It was a mutual loss, for both Schwinger and the community were the losers. – Silvan Schweber (S1984, p. 371-2)**

The invisible quark. So far I have said little about quarks or gluons, except to mention that the strong pion field is made of more basic fields called quarks and gluons. Now I must tell you that the baryon field also is made of quarks and gluons. All other fields — gravity, EM, weak, and lepton — are basic, but the baryon and strong fields are made of more fundamental fields. We will say more about this in Chapter 10, but in the meantime, please don't worry about it. It is a complication that does not interfere with your understanding the essence of QFT.

Note about youth. Before ending this chapter, I'd like to say something about youth. It was a popular edict in the 1960's never to trust anyone over 30. Whether or not this is true in general, it has great validity in theoretical physics. When Einstein discovered his theory of relativity he was 25. Bohr was 28 when he developed his atomic theory. De Broglie was 28 when he had his epiphany. Pauli was 25 when he announced his exclusion principle. Heisenberg was 26 when he presented his uncertainty principle (although the exact date is uncertain). Dirac was 26 when he discovered his relativistic equation. Newton was 23 when he worked out his gravitational theory. Even

Schwinger and Feynman (but not Tomonaga), despite the interruption of WWII, found the renormalization solution before the age of 30. It seems that the ability to “think outside the box” is strongest in the young!

Age is, of course, a fever chill
That every physicist must fear.
He’s better dead than living still
When once he’s past his thirtieth year.
— P. Dirac (A2004, p. 254)

SUMMARY

We have now added two more quantum fields to our palette, bringing the final total to seven. These matter fields differ from force fields in two important ways. First, they do not exert forces; they create force fields and are acted on by force fields. Second, their spin is ½ (in Planck units). As we will show in the next chapter, this half-integer value leads to the Pauli Exclusion Principle which states that only one field quantum can occupy a given state at a time. This explains why matter quanta behave much more like particles.

There are two types of matter fields: leptons (electrons, muons, neutrinos, etc.) and baryons (neutrons, protons, etc.). In our imaginary color scheme they are represented by yellow and red respectively. The baryon fields have a greater mass (hence the name baryon), and are the only fields that interact with the strong force field. The complete table of fields in QFT is as follows:

Field	Spin	Quantum	Interacts with	"Color"
Gravity	2	Graviton	everything	Blue
EM	1	Photon (γ)	charged fields	Green
Strong*	0	Pion (π^+, π^-, π°)	nucleon fields	Purple
Weak	1	(W^+, W^-, Z°)	leptons &baryons	Brown
Higgs	0	Higgs boson	everything	?
Matter – leptons	½	Electron (e), Muon (μ), Neutrino (ν), etc.	N/A	Yellow
Matter – baryons*	½	Proton (p) Neutron (n), etc.	N/A	Red

*We will see in Chapter 10 that the strong field and the baryon field are made of more basic but "invisible" fields called quarks and gluons.

CHAPTER 7
SPECIAL RELATIVITY

> **A while back I bought an idiot's guide to relativity, from which I learned only that I hadn't yet attained the rank of idiot. – Joe Bennett (The Press, Christchurch, July 4, 2007)**

In 1905 Einstein published "On the electrodynamics of moving bodies", the third of his four *annus mirabilis* (miracle year) papers. In this paper he introduced his famous *Principle of Relativity*: "The same laws of electrodynamics and optics will be valid for all frames of reference for which the equations of mechanics hold good". In other words, things behave the same way regardless of one's state of motion, as long as it's uniform. From this simple principle there follow some surprising results, such as:

- Objects contract when they're moving.
- Time slows down for objects (and people) in motion.
- Mass increases with motion.
- Nothing can go faster than light.

Let's face it, to most people these behaviors don't make sense. Why should such things happen? We will show that in a world made of fields, not particles, these statements are a natural consequence of the way fields behave. You will see that they are not only understandable, but inevitable. And along the way we will clear up a few misconceptions.

Bottom-up method. Using the behavior of fields to explain these results is what I call the *bottom-up* approach, as contrasted with Einstein's *top-down* method of deriving them from the Principle of Relativity. Physicists today prefer the top-down method because it makes calculations easier, but the bottom-up approach provides an explanation for what otherwise seem to be paradoxical effects. The Principle of Relativity itself is a consequence of the field equations. As John Bell[1] wrote:

> **Einstein starts from the *hypothesis* that the laws will look the same to all observers in uniform motion. This permits a very concise and elegant formulation of the theory, as often happens when one big**

assumption can be made to cover several less big ones… But in my opinion there is also something to be said for taking students along the road made by FitzGerald, Larmor, Lorentz and Poincaré. The longer road sometimes gives more familiarity with the country. — J. Bell (B2001, p. 70)

The beginning. The story of relativity did not begin in 1905. It started in 1881 with an experiment that produced surprising results — results that helped lead Einstein to his theory.[2] Shortly before his death, James Maxwell had proposed that the earth's motion through the ether (which was still believed in at the time) could be determined by measuring the speed of light parallel to the earth's motion and perpendicular to that motion. By comparing these two measurements, he said, one could calculate the earth's speed as it passes through the ether. However, the measurement accuracy that would be needed (one part in 200 million) was well beyond the capability of the time, so Maxwell concluded that the experiment was impossible. It took a young American physicist to make it possible.

Albert Michelson (1852-1931). Michelson came to the United States at the age of two, the son of Jewish-German parents. After serving in the US Navy (which he rejoined at the age of 62 to serve in World War I), he pursued a career in physics. In 1881, while studying in Europe, he came across Maxwell's "challenge" and conceived the idea of the interferometer – an instrument that can measure exceedingly small distances by observing optical interference patterns. Using this instrument, Michelson was able to perform Maxwell's experiment, and the result caused a major upheaval in physics that culminated in Einstein's theory of relativity. Yet when Michelson received the 1907 Nobel prize in physics, this experiment was ignored, because the theory of relativity was too controversial at the time. (The same thing happened to Einstein, as we will see.) Instead the award was given for "optical precision instruments and the spectroscopic and metrological investigations carried out with their aid".

The apparatus. The central part of Michelson's apparatus is a thinly-silvered mirror that splits a light beam into two parts, with one beam traveling through the mirror and the other reflected upward. The two beams are then reflected back to the central mirror, which sends them to a detector. If the apparatus is stationary and if the two paths are equal in length, then the light beams would take equal times to reach the detector (Fig. 7-1). Since it was not possible to make the two lengths exactly equal, Michelson arranged to rotate the apparatus 90° so that first one beam and then the other would lie in the direction of the earth's motion, allowing him to correct for the difference.

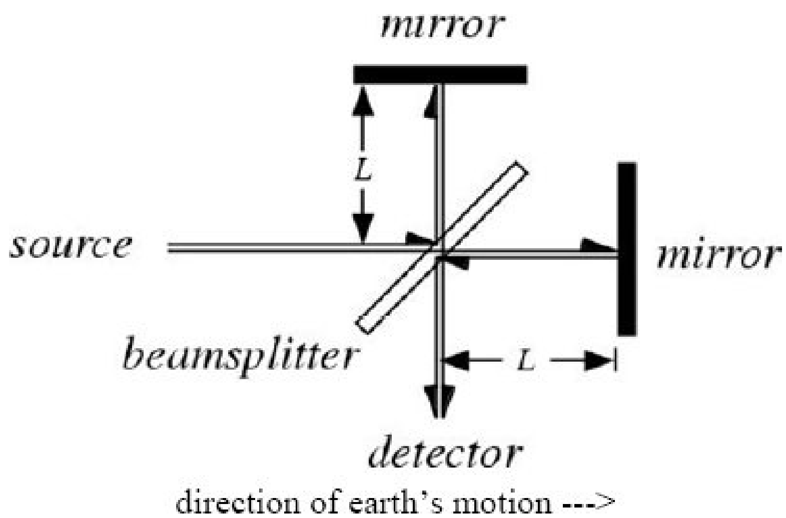


Fig. 7-1. Michelson's apparatus. The path lengths between the beam splitter (center) and the two mirrors are equal but the beam moving to the right must travel a greater distance because of the earth's motion.[3]

If the apparatus is moving, however, the situation is different. Because the mirrors are moving, the light beam traveling in the direction of motion must cover a greater distance. (The transverse beam would also be affected by motion, but not as much.[4]) Then when the two beams are recombined at the detector, they will create an interference pattern similar to that seen in the two-slit experiment (Fig. 3-3b).

The result. When the experiment was performed, much to Michelson's surprise, the two light beams took the same time to reach the detector despite the extra distance that one of them had to travel! More accurate experiments were performed later in collaboration with Edward Morley, using more mirrors to extend the path lengths. This improvement in accuracy turned out to be critical, as Michelson had made an error in his first measurement that was pointed out by Hendrik Lorentz. The experiment, now called the Michelson-Morley (M-M) experiment, was repeated at different times of day (as the earth's surface moves in different directions because of its rotation) and at different seasons of the year (when the earth moves in different directions as it orbits the sun). The answer remained the same: The two light beams took equal times to traverse their paths, regardless of the earth's motion.

A surprise. That the speed of light should be independent of motion was most surprising. Einstein even felt the need to make it a separate postulate: "light is always propagated in empty space with a definite velocity which is independent of the state of motion of the emitting body", adding that this is "only apparently irreconcilable with the [Principle of Relativity]".[5]

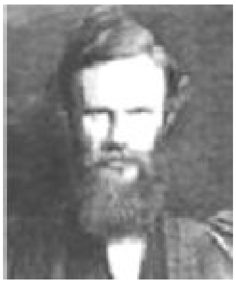
The reader should pause to appreciate how strange is this statement. It makes no sense for a light beam – or anything, for that matter – to travel at the same speed regardless of the motion of the observer. Suppose, for example, that you are observing a very fast train from another train. The apparent speed of the fast train would clearly depend on its direction relative to yours. If the other train is moving in the opposite direction, it would go whooshing by (Fig. 7-2), but if it is moving in the same direction, it would pass very slowly. Yet Michelson, a passenger on a train called earth, found that another train called light always moves at the same speed no matter which way earth is moving.

Fig. 7-2. A passenger in a train would be surprised if a passing train always moved at the same speed, regardless of the direction of its motion.

If the M-M experiment had been performed only once, there would have been no problem. We could have simply said that this is the frame of reference in which the laws of physics hold — in which Maxwell's equations apply and light travels with velocity c. But the result was always the same, regardless of earth's motion. It is not possible for light to travel with the same velocity in all of these frames of reference unless "something funny" is going on.

LENGTHS CONTRACT

The "something funny" turned out to be even more surprising than the M-M result itself: Objects contract when they move! More specifically, they contract in the direction of motion. Think about it. If the path length of Michelson's apparatus in the forward direction contracts by the same amount as the extra distance the light beam would have to travel because of motion, the two effects would cancel out. In fact, this is the only way that Michelson's null result could be explained. The idea of contraction was first suggested by a relatively unknown Irish physicist.



George Francis FitzGerald 1851-1901. FitzGerald expressed his idea in a short communication to the American journal Science in 1891, ten years after Michelson's first reported result. He also suggested a reason.

> **I would suggest that almost the only hypothesis that can reconcile this [conflict] is that the length of material bodies changes, according as they are moving through the ether or across it, by an amount depending on the square of the ratio of their velocities to that of light. We know that electric forces are affected by the motion of the electrified bodies relative to the ether, and it seems a not improbable supposition that the molecular forces are affected by the motion, and that the size of a body alters consequently. — G.F. FitzGerald (P1982, p. 122)**

While FitzGerald referred to the ether, which was believed to be the carrier for light waves at the time, the reasoning holds with or without the ether. Somewhat later his rather timid suggestion that molecular forces are affected by motion was repeated and refined by the most famous physicist of the time.

Hendrik Lorentz (1853-1928). While FitzGerald was little known outside Ireland, the Dutch scientist Hendrik Lorentz was recognized as the greatest physicist since Maxwell. In 1902 he and Pieter Zeeman received the second Nobel Prize ever awarded for discovering the *Zeeman effect* that, as you may recall, led to the discovery of electron spin (see Chapter 6). To Einstein, Lorentz was“the most well-rounded and harmonious personality he had met in his entire life” (P1982, p. 169). Upon Lorentz’s death, Europe’s greatest physicists attended his funeral and three minutes of silence were observed throughout Holland.

Lorentz had not seen FitzGerald’s paper, but he too realized that Michelson’s strange result would make sense if the apparatus contracted along the direction of motion. However, he went further than FitzGerald; he did the calculation (not an easy one). When he found that the theoretical contraction *exactly* compensated for the extra travel distance, this was one of the great “eureka” moments in physics, comparable to those of Newton and Einstein.

> **Now it is very remarkable that we find exactly the amount of change that was postulated, if we extend to molecular actions the result found for the electric forces. — H. A. Lorentz (L1915, p. 202)**

When Lorentz learned of FitzGerald’s work, he wrote to him to be sure he was not usurping credit. FitzGerald, a modest man, wrote back as follows:

> **As I recollect, I wrote a letter to Science… but I do not know if they ever published it as I did not see the journal for some time afterwards. I am pretty sure that your publication was then prior to any of mine. I am particularly delighted that you agree with me for I have been rather laughed at for my view over here… But now that I have you as an advocate and authority I shall begin to jeer at others for holding any other view. – G. F. FitzGerald (B2001, p. 236)**

FitzGerald’s paper *had* been published, and thereafter Lorentz was careful to acknowledge FitzGerald's priority. Today the effect may be called the Lorentz contraction, the FitzGerald-Lorentz contraction, or the FitzGerald contraction. Fig. 7-3 shows a modern version of Lorentz's calculation done by John Bell with the aid of a computer.

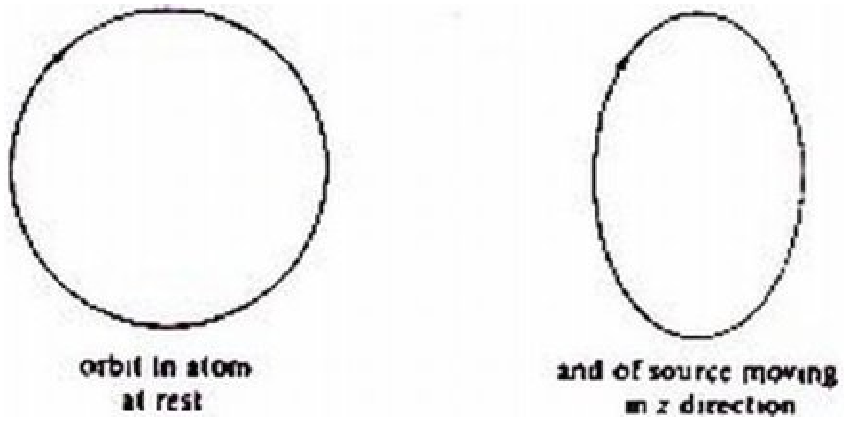


Fig. 7-3. John Bell's calculation of an electron orbit (regarding the electron as a particle) both at rest and in motion. If Bell had calculated the shape of the electron field (Fig. 6-2) instead of particle orbits, it would show the same foreshortening. (B2003, p. 64).

Misconception #1. Some writers claim that the Lorentz contraction was an ad hoc explanation offered without any theoretical basis. In fact, it was based on a deep understanding of how fields behave when in motion and how this behavior affects the molecular configurations.

> **Surprising as this hypothesis may appear at first sight, yet we shall have to admit that it is by no means far-fetched as soon as we assume that molecular forces are also transmitted through the ether, like the electric and magnetic forces of which we are able at the present time to make this assertion definitely. If they are so transmitted, the translation will very probably affect this action between two molecules or atoms in a manner resembling the attraction or repulsion between charged particles. Now, since the form and dimensions of a solid body are ultimately conditioned by the intensity of molecular actions, there cannot fail to be a change of dimensions as well. — H. A. Lorentz (E1923, p. 5-6)**

Everything contracts. It is important to understand that it was not just Michelson's apparatus that contracted, it was anything and everything on earth, including Michelson himself. As John Bell wrote about a moving observer:

> **But will she not *see* that her meter sticks are contracted when laid out in the [direction of motion] – and even decontract when turned in the [other] direction? No, because the retina of her eye will also be contracted, so that just the same cells receive the image of the meter stick as if both stick and observer were at rest. – J. Bell (B2001, p. 68)**

Furthermore, it is not just objects. Space also contracts because space, as we have seen, is not empty. It is filled with fields, and these fields are always interacting with each other, even if there are no atoms or molecules present. So not only do objects contract, space itself contracts. Anything made of fields contracts when in motion.

Relative to what? If the contraction of Michelson's apparatus was invisible to him, one may well ask, to whom was it visible? The answer is, to a stationary observer not on the earth. She (continuing Bell's use of the feminine pronoun) would see that one arm of the apparatus is shorter, and she would understand immediately why Michelson, who was unaware of the contraction, found no difference in travel time between the two directions.

Misconception #2. It is sometimes said that the M-M experiment disproved the ether hypothesis. This is not true. The Lorentz contraction is a consequence of Maxwell's Equations, with or without the ether. The fact that both FitzGerald and Lorentz believed in the ether is irrelevant. In QFT the fields have a reality in themselves and can be thought of as a kind of ether (see "Goodbye ether" in Chapter 3):

> **More careful reflection teaches us, however, that the special theory of relativity does not compel us to deny ether… To deny the ether is ultimately to assume that empty space has no physical qualities whatever. — A. Einstein (I2007, p. 318)**

> **Moreover space-time itself had become a dynamical medium – an ether, if there ever was one. — F. Wilczek ("The persistence of ether", Physics Today, Jan. 1999, p. 11).**

Bell's breaking thread. To illustrate the reality of the Lorentz contraction, John Bell proposed the following thought experiment. Suppose that two rocketships, connected by a tightly-stretched thread, take off from earth simultaneously, one behind the other. Suppose further that during the takeoff and acceleration they maintain exactly the same separation as observed from earth. That is, if they were 100 meters apart at takeoff, they remain 100 meters apart even after reaching final speed. The question is, what happens to the thread? The answer is that it breaks because of the Lorentz contraction. Surprisingly, even some physicists find this hard to accept.

> **A distinguished experimental physicist refused to accept that the thread would break, and regarded my assertion, that indeed it would, as a personal misinterpretation of special relativity… Many people who give this wrong answer at first get the right answer on further reflection. Usually they feel obliged to work out how things look to observers [on the rocketships]… It is only after working this out, and perhaps only with a residual feeling of unease, that such people finally accept a conclusion which is perfectly trivial in terms of the account of things [from the earth], including the Fitzgerald contraction. – J. Bell (B2001, p. 62)**

The Ehrenfest paradox. Another challenging example of Lorentz contraction is the *Ehrenfest paradox,* proposed by Paul Ehrenfest in 1909. He asked what will happen if a rigid disk rotates so that the rim is traveling at a speed close to that of light. Will the circumference contract, and if so won't that violate the geometrical relation between the circumference of a circle and its diameter? Einstein's answer, in his 1916 paper "The Foundation of the General Theory of Relativity", was:

> **…the quotient [of the circumference and diameter] would be greater than π. This is readily understood if we envisage the whole process of measuring from the "stationary system", and take into consideration that the measuring rod applied to the periphery undergoes a Lorentz contraction, while the one applied along the radius does not. Hence Euclidean geometry does not apply…" – A. Einstein (E1923, p. 116)**

However, Einstein didn't say what would happen to the disk itself, and the question is controversial even today.[6]

I think there is a simple answer. Consider a merry-go-round as it begins to rotate, and imagine that you are riding a horse at the rim. Now consider Bell's "breaking thread" experiment, expanded to a million rocketships circling the earth, each connected by a thread to the one in front. As Bell showed, each string will break during the acceleration because of the Lorentz contraction. Now if we consider the horses as rocketships circling the center of the merry-go-round, it is clear that the material connecting the horses will fracture as the merry-go-round speeds up. In fact, it is not just the area between horses where this will happen, the entire rim (horses included) will break into small segments, with cracks between them.

Note that this contraction effect is different from the centrifugal force that also acts on the merry-go-round. That force can be resisted if the interatomic binding is strong enough, but the Lorentz contraction, cannot be resisted. It is caused by actual shrinkage of the atoms (see Fig. 7-3) that necessarily disrupts the interatomic binding. So the answer to Ehrenfest's paradox is that the disk will develop cracks or gaps, no matter how "rigid" it is, and break up into many pieces with distorted shapes.

Now let us suppose that you somehow manage to remain in one piece and get off your horse to measure the circumference. You do this by laying down a measuring stick repeatedly, being careful to step over the gaps. Because the stick has contracted, it will take more stick-lengths to cover the circumference than when the merry-go-round was at rest. You may therefore conclude, as Einstein did, that "Euclidean geometry does not apply." However, there is another possibility. Being aware of the strong centrifugal force, you might conclude — especially if you are a physicist — that the merry-go-round is rotating and the meter-stick has undergone Lorentz contraction. Either way there is no paradox.

TIME DILATES

There once was a spaceman named Dwight
Who could travel much faster than light.
He departed one day, in a relative way
And returned on the previous night.

This, of course, is a joke. No one can travel faster than light or go backward in time, but it *is* true that time passes more slowly for people or objects that are moving. When I say time passes more slowly, I mean that things take longer to happen. Fruit doesn't ripen as fast, beards take longer to grow, and clocks run more slowly. This is known as *time dilation*. Of course the effect is exceedingly small at ordinary speeds, but it becomes significant if the speed approaches that of light. As with length contraction, the slowdown is not noticeable in the moving system:

> **She will not notice that her clocks have slowed down, because she will herself be thinking more slowly. – J. Bell, 1976 (B2001, p. 68)**

The twin paradox. Time dilation is probably the best-known of the relativity effects because of the *twin paradox*. Here is the scenario: An astronaut leaves on a rocketship traveling at close to the speed of light. After whizzing around the galaxy he returns to find that his twin brother on earth is an old man with a long beard while he himself is still young. Now this is certainly mind-boggling. Why should time pass more slowly just because you're moving? What physical explanation can we find for that? In the case of contraction, I asked you to accept on faith that it follows as a natural consequence of the field equations. In the case of time dilation (and the remaining paradoxes) I will provide an intuitive explanation based on the way fields behave.

Intuitive explanation. Consider two atoms in a rocketship. Suppose that the rearward atom creates a field disturbance and when that disturbance reaches a more forward atom, something happens. (It is the interaction among atoms, after all, that causes things to happen.) By the time the disturbance reaches the second atom it will have moved farther ahead, so the disturbance must travel a greater distance than if the rocketship were stationary (even after taking the Lorentz contraction into account). Since fields propagate at a fixed rate, it will take a longer time (as observed from the earth) for the disturbance to reach the second atom.[7] In short, things happen more slowly when you're moving because the fields have to travel a greater distance.

An analogy. Consider two men on a raft calling back and forth to each other (Fig. 7-4). Suppose that when each man receives the information he causes certain things to happen. The problem is that it takes time for the sound waves to travel, so by the time the sound from A reaches B, B will have moved to position B', and the communication time will be longer. The reverse is true when B calls back to A, but this decrease in travel time isn't as great. Another complication is that the space between the two men contracts, making the calculation difficult. However, if the men are aligned perpendicular to the motion, as in the upper picture, the calculation is easier, and can be done using high school algebra. It is not hard to show that the result is the same as Lorentz's result.[8]

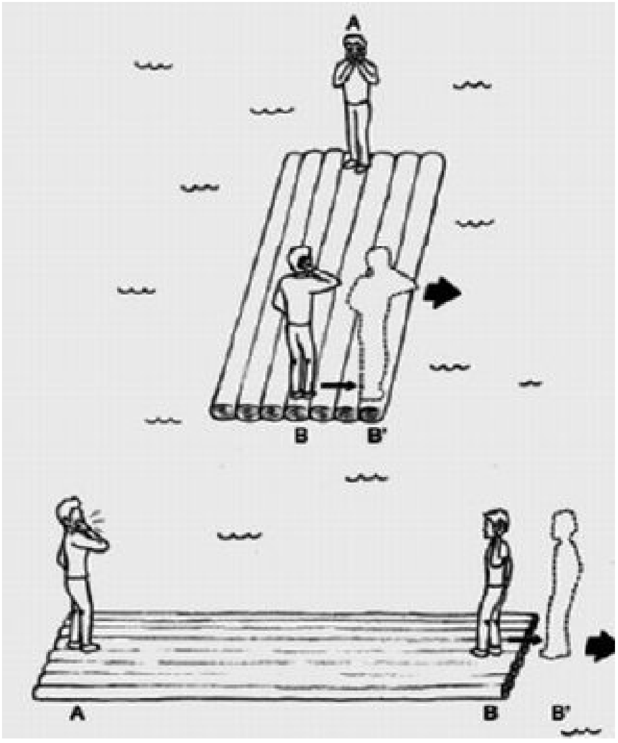


Fig. 7-4. Two men standing on rafts moving to the right. In the top picture they are aligned perpendicular to the motion, and in the bottom picture they are aligned parallel to the motion.

Misconception #3. Some physicists claim that length contraction and time dilation are not real and that the physical explanations of FitzGerald, Larmor and Lorentz are not to be taken seriously. This is not true.

> **Moving clocks really do run slowly and moving sticks really do shrink … It is necessary for clocks and sticks really so to behave if the whole subject is to fit coherently together, and not collapse into a mass of self-contradiction. — N. David Mermin (M2005, Chap.**

13)

NASA routinely observes time dilation in orbiting satellites and corrections are applied to keep atomic clocks on the GPS satellites in sync with clocks on earth. Time dilation has also been seen in particle accelerators. At the CERN accelerator radioactive particles traveling at 99.9% the speed of light decay 30 times more slowly than they do at rest (S1986, p. 57).

Another analogy. The idea of length contraction and time dilation may be easier to accept when you consider that objects contract and processes slow down at low temperatures. As Lorentz said:

> **We may, I think, even go so far as to say that… the conclusion is no less legitimate than the inferences concerning the dilatation by heat. - H. Lorentz (L1916, p. 196**

Would we think it paradoxical if a twin was placed in a cold chamber for 50 years and then emerged to find that his brother was old and had grown a long beard, while he remained young? No, we would not; in fact there are firms that offer to preserve people by freezing them. Why then should we not accept that motion can have a similar effect on chemical and physical processes?

THE SPEED LIMIT

Of course the idea that there is an ultimate speed limit seems absurd. While the speed of light is very high by earthly standards, the magnitude is not the point; any kind of speed limit in nature doesn't make sense. Suppose, for example, that a spaceship is traveling at almost the speed of light. Why can't you fire the engine again and make it go faster – or if necessary, build another ship with a more powerful engine? Or if a proton is whirling around in a cyclotron at close to the speed of light, why can't you give it additional energy boosts and make it go faster?

Intuitive explanation. When we think of the spaceship and the proton as made of fields, not as solid objects, the idea is no longer ridiculous. Fields can't move infinitely fast. Changes in a field propagate in a "laborious" manner, with a change in intensity at one point causing a change at nearby points, in accordance with the field equations. Consider the wave created when you drop a stone in water: The stone generates a disturbance that moves outward as the water level at one point affects the level at another point, and there is nothing we can do to speed it up. Or consider a sound wave traveling through air: The disturbance in air pressure propagates as the pressure at one point affects the pressure at an adjacent point, and we can't do anything to speed it up. In both cases the propagation speed is determined by properties of the transmitting medium — air and water — and there are mathematical equations that describe those properties.

Fields are also described by mathematical equations, based on the properties of space, and the constant *c* in those equations determines the maximum speed of propagation. If the field has mass there is also a mass term that slows the speed. Since everything is made of fields — including protons and rocketships — it is clear that nothing can go faster than light. As Frank Wilczek wrote,

> **One of the most basic results of special relativity, that the speed of light is a limiting velocity for the propagation of any physical influence, makes the field concept almost inevitable. – F. Wilczek ("The persistence of Ether", p. 11, Physics Today, Jan. 1999)**

David Bodanis tried to make this point in the following way:

> **Light will always be a quick leapfrogging of electricity out from**

> **magnetism, and then of magnetism leaping out from electricity, all swiftly shooting away from anything trying to catch up to it. That's why its speed can be an upper limit. – D. Bodanis (B2000, p. 50)**

However, Bodanis only told part of the story. It is only when we recognize that everything — not just light — is made of fields that we can conclude that there is a universal speed limit.

Now let's take another look at that proton whirling around in an accelerator, using our colored glasses to visualize the fields. We see the proton as a blob of redness oozing (I prefer that term to "leapfrogging") ahead, as the amount of redness at one point affects the redness at a neighboring point. The process is very fast by our usual norms, but it is not instantaneous. The proton can't move any faster because the field equations put a limit on how fast the redness can ooze.

MASS INCREASES

The final paradox of relativity is the increase in mass due to motion. Mass increase has been observed experimentally in particle accelerators, with increases as great as 3000% for particles traveling at over 99.9% the speed of light. How can the mass of an object get bigger just because it is moving?

Intuitive explanation. As we saw in Chapter 2, mass means inertia, i.e., resistance to acceleration. If you push something and it doesn't respond much, then by definition it has a large mass or inertia. Now we just saw that pushing on something that is traveling at close to the speed of light has little effect on its speed because the underlying fields are already moving almost as fast as they can. Thus its resistance to acceleration has become greater and this means its mass has increased. Mass increase is just another way of saying that fields can't propagate faster than c.

RASHOMON REALITY

The confusing thing about all this is not so much the relativistic effects themselves, which make eminent sense from the field point of view, but that reality is seen differently by different observers. In the above examples, the reality described is what a stationary observer would see, but observers in the moving system would see it very differently. Time dilation, Lorentz contraction, and mass increase are not perceived by the observers in motion. Even Einstein found this confusing:

> **The question of whether the Lorentz contraction does or does not exist is confusing. It does not 'really' exist in so far as it does not exist for an observer who moves… it really exists, however, in the sense that it can as a matter of principle be demonstrated by a resting observer. — A. Einstein 1911 (P1982, p. 144)**

The idea that a single reality can be perceived differently by different observers was the basis for the Japanese movie Rashomon, so I call it *Rashomon reality*. Of course it is not psychological factors that cause observers to see things differently; it is different states of motion. If different observers have measuring rods (and eyeballs) of different lengths, and clocks (and brains) that run at different speeds, of course they are going to see things differently. Events that are simultaneous to one observer may not be simultaneous to another and things fast become confusing — even for physicists.

Now let us look again at the twin paradox. Since motion is relative, each twin must believe that it is the other's clock (and body) that has slowed down. How, then, can it be that when they reunite, the spaceman has aged less than the man on earth? The answer is not easy to understand, but for those who have the stomach for it, Schwinger's explanation is given in a footnote.[9] A similar situation exists in the breaking thread experiment. While observers on earth see the thread break because of Lorentz contraction, observers on the rocketships don't see any contraction. They see the ships drifting apart as they accelerate, and this is what causes the thread to break.

I recommend that lay readers leave these details to the physicists and simply accept that the same reality is seen differently by different observers. Rashomon reality is tricky, but all you need to know is that time and dimensions are affected by motion. Yes, there is a common reality, and it is made of fields, but it is perceived differently by observers in different states of motion.

TOP-DOWN vs. BOTTOM-UP

Most physicists today are "top-downers". Following Einstein, they regard the relativistic effects as consequences of the Principle of Relativity. They believe that to explain these effects as consequences of the way fields behave is somehow illegitimate. "Look what you've done to our beautiful theory," they say. "You've reduced it to mere physical effects. The Lorentz contraction is not a physical process that occurs because field configurations are affected by motion; it is something that is built into the nature of space. And this time dilation — it's not that processes happen more slowly, it is a property of time itself."

Not surprisingly, Einstein preferred the top-down approach:

> **Lorentz and FitzGerald rescued the theory… by assuming that the motion of the body relative to the ether produces a contraction of the body… from the standpoint of the theory of relativity this solution… was the right one. But on the basis of the theory of relativity the method of interpretation is incomparably more satisfactory… Here the contraction of moving bodies follows from the two fundamental principles of the theory, without the introduction of particular hypotheses. – A. Einstein (E1961, p. 59)[10]**

Lorentz, not surprisingly, preferred bottom-up:

> **Einstein simply postulates what we have deduced, with some difficulty and not altogether satisfactorily, from the fundamental equations of the electromagnetic field. By doing so, he may certainly take credit for making us see in the negative result of experiments… not a fortuitous compensation of opposing effects, but the manifestation of a general and fundamental principle. Yet, I think, something may also be claimed in favor of the form in which I have presented the theory. – H. Lorentz (L1916, p. 229-230)**

The fact is, either approach is correct, and one does not preclude the other. Yes, the Principle of Relativity is elegant and the top-down approach is easier to use; physicists love it for that reason. But the field equations are also elegant and they not only contain the Principle of Relativity within them, they also provide a physical explanation for effects that otherwise seem paradoxical. We can never know if God started with the Principle of Relativity and derived the field equations or started with the field equations from which follows the Principle. If She started with the principle that the laws of nature should be the same in all moving systems, then She also provided mechanisms to make it happen. And if the mechanisms are there, why not use them? They are real and understandable, and they should not be ignored. Here is how John Bell summed it up:

> **The facts of physics do not oblige us to accept one philosophy rather than the other… the laws of physics in any one reference frame account for all physical phenomena, including the observations of moving observers. And it is often simplest to work in a single frame, rather than to hurry after each moving object in turn… It is a great pity that students don't understand this. Very often they are led to believe that Einstein somehow swept away all that went before. This is not true. – J. Bell (B2003, p. 70, 238)**

Karen vs. Albert. I would like to conclude this chapter by quoting my wife, Karen, who came to this book with little knowledge of physics and even less interest (or was it the other way around?), but who has been of great help in its writing. When I asked her what she thought about time dilation and spatial contraction being "properties of space and time", she said something like this:

> **How can something which is no more than a marker of events in itself have special properties? If there are any special properties, it is in the way those events proceed, not in something as abstract as "time". The same is true of space. Space has no meaning in itself; only the word "distance" does, and distance has meaning only in relationship to physical objects.**

I was so impressed that I rushed to write it down. Then a few days later I encountered the following words by Einstein:

> **Space-time is not necessarily something to which one can ascribe a separate existence, independently of the actual objects of physical reality. Physical objects are not *in space*, but these objects are *spatially extended*. In this way the concept "empty space" loses its meaning. — A. Einstein (E1961, p. vii)**

Am I wrong in thinking that Karen said it better?[11] In any case, I hope that Karen's insight will encourage other lay readers to work at understanding these concepts. Science is too wonderful and rewarding an activity to be left only to the scientists.

SUMMARY

The theory of Special Relativity, formulated by Einstein in 1905, is based on the postulate that the laws of physics are the same regardless of the state of motion of the observer, so long as it is uniform. This is known as the Principle of Relativity, from which there follow many strange effects. While these behaviors seem paradoxical, they make perfectly good sense when seen as a result of the way fields behave, whether they be quantum fields or classical fields.

- Objects contract when moving because motion affects the interaction of fields that hold the object together. Space itself contracts because space is made of fields.
- Things happen more slowly in a moving system because the interacting fields must travel a greater distance (despite the contraction).
- Nothing can go faster than light because everything is made of fields that propagate at a finite rate determined by the field equations.
- Mass increases with speed because mass means resistance to acceleration and acceleration beyond the speed of light is impossible.

I call this the *bottom-up* approach. Although most physicists prefer to start with the Principle of Relativity, the bottom-up method provides insight into why these strange things happen. Even the Principle of Relativity follows from the bottom-up approach.

Either way, one must cope with a Rashomon reality in which observers in differently-moving systems see the same reality in different ways. Rashomon reality is very confusing, and is best left to the physicists to deal with.

CHAPTER 8
GENERAL RELATIVITY

> **When a journalist asked the British astronomer Sir Arthur Eddington if it was true that he was one of only three people in the world who could understand Einstein's relativity theories, Eddington considered deeply for a moment and replied: "I am trying to think who the third person is." – B. Bryson (B2003, p. 124)**

General Relativity is the name Einstein gave to his theory of gravity. As the theory is usually presented, gravity is said to be caused by curvature in four-dimensional space-time. This is a concept beyond the reach of ordinary folks like you and me.

> **The non-mathematician is seized by a mysterious shuddering when he hears of 'four-dimensional' things, by a feeling not unlike that awakened by thoughts of the occult. — A. Einstein (E1961, p. 61)**

In this chapter we will see that in QFT (and also in Einstein's theory) there is no eerie fourth dimension: space is space and time is time. We will also see that gravity is caused by a force field — not curvature, and that, contrary to popular belief, QFT is compatible with general relativity.

SPACE-TIME ISN'T 4-DIMENSIONAL

> **The most challenging and non-intuitive of all the concepts in the general theory of relativity is the idea that time is part of space… Our brains can take us only so far because it is so nearly impossible to envision a dimension comprising three parts space to one part time, all interwoven like the threads in a plaid fabric. – B. Bryson (B2003, p. 125-126)**

Of course space-time is four-dimensional in the trivial sense that it takes four numbers to specify when and where an event takes place, but that doesn't mean that space and time are equivalent. In QFT, as in Einstein's theory, space and time play separate roles in accord with our natural perceptions. The idea that space-time must be viewed as a four-dimensional entity was introduced by the German mathematician Hermann Minkowski. In a speech at Cologne in 1908, he expressed this view with great eloquence:

> **Henceforth space by itself, and time by itself are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality… It is only in four dimensions that the relations here taken under consideration reveal their inner being in full simplicity, and that on a three-dimensional space forced upon us *a priori* they cast only a very complicated projection. — H. Minkowski (E1923, p. 75, 90)**

Einstein, however, did not subscribe to that view; he called it "superfluous erudition". When adapting Minkowski's mathematical formalism to describe the gravitational field, Einstein had to add the "imaginary" number *i*[1] to the time term. This makes the equations more compact and easier to work with, but Einstein was careful to distinguish the formal aspect of the notation from reality (emphases added):

> **The discovery of Minkowski… was of importance for the *formal* development of the theory of relativity… The four-dimensional space-time continuum of the theory of relativity, in its most essential *formal* properties, shows a pronounced relationship to the three-dimensional continuum of Euclidean geometrical space… Under these conditions, the natural laws… assume *mathematical forms* in which the time coordinate plays exactly the same role as the three space coordinates. — A. Einstein (E1961, p. 63)**

Einstein is saying that the four-dimensional notation is useful for physicists; it is a convenient way of handling the mathematical relationship between space evolution and time evolution that is required by special relativity. One might almost say that physicists couldn't live without it. Nevertheless, space and time are different, and I say shame on those who try to foist and force the four-dimensional concept onto the public as essential to the understanding of relativity theory.

GRAVITY ISN'T CURVATURE

In most presentations of physics today we are told that gravity is caused by "curvature of space-time". This was not Einstein's view, nor is it the view of QFT. Einstein believed that gravity is a force field, not unlike the electromagnetic field.

> **[There is] a field of force, namely the gravitational field, which possesses the remarkable property of imparting the same acceleration to all bodies. – Albert Einstein (E1923, p. 114)**

The idea of space-time curvature, like the four-dimensional concept, had its origin in mathematics. When searching for a mathematical method that could embody his Principle of Equivalence, Einstein was led to the equations of Riemannian geometry. And yes, these equations describe four-dimensional curvature, for those who can visualize it. Mathematicians are not limited by physical constraints; equations that have a physical meaning in three dimensions can be generalized algebraically to any number of dimensions. But when you do this, you are dealing with algebra, not geometry.

> **To those who are geometrically inclined, two dimensions is a breeze, three dimensions routine, and four dimensions impossible. But to those who think algebraically, two, three, or four dimensions are just particular examples of spaces with any number of dimensions. In this sense, Riemann was an algebraist. – J. Schwinger (S1986, p. 175-176)**

Because this is such a controversial question, I will quote two more Nobel laureates who expressed similar thoughts:

> **We can describe general relativity using either of two mathematically equivalent ideas: curved space-time or metric field. Mathematicians, mystics, and specialists in general relativity tend to like the geometric view because of its elegance. Physicists trained in the more empirical tradition of high-energy physics and quantum field theory tend to prefer the field view… More important, as we'll see in a moment, the field view makes Einstein's theory of gravity look more like the other successful theories of fundamental physics, and so makes it easier to work toward a fully integrated, unified description of all the laws. As you**

can probably tell, I'm a field man. — F. Wilczek (W2008, p. 100-101)

It is certainly a historical fact that when Albert Einstein was working out general relativity, there was at hand a preexisting mathematical formalism, that of Riemannian geometry, that he could and did take over whole. However, this historical fact does not mean that the essence of general relativity necessarily consists in the application of Riemannian geometry to physical space and time. In my view, it is much more useful to regard general relativity above all as a theory of gravitation, whose connection with geometry arises from the peculiar empirical properties of gravitation. – S. Weinberg (W1972, p. vii, p.3)

A physicist friend of mine put it more succinctly: "Why would God invent a different mechanism for another force?"

GRAVITY AND QFT ARE COMPATIBLE

It is often said that general relativity is incompatible with quantum theory. Julian Schwinger did not agree.

> **[Consider] a neutral field that presumably possesses no internal properties and responds dynamically to the space-time attributes of other systems… It appears that in the hierarchy of fields there is a natural place for the gravitational field. — J. Schwinger (S1957, p. 433)**

Schwinger went on to publish two papers on “The quantized gravitational field” in the Physical Review in 1963.

This is not to say there is no problem with quantum gravity. Just as the QFT equations for the EM field led to infinite values, so the gravitational field equations lead to infinities, but these infinities cannot be circumvented by renormalization, as described in Chapter 6. But this does not mean that QFT and general relativity are inconsistent. It only means that the interaction of a gravitational quantum with its self-field is not described by the theory (see “The gaps” in Chapter 10).

Although renormalization doesn’t work for quantum gravity, Schwinger found another way around the problem of the infinities, using a method he called *source theory*. Using this method, he was able to reproduce all four of Einstein’s classic results: gravitational red shift, deflection and slowing down of light by gravity, and the perihelion precession of Mercury (S1970, p. 82-85). The neglect of source theory by the physics community was a major disappointment for Schwinger:

> **The lack of appreciation of these facts by others was depressing, but understandable — J. Schwinger (S1970, Preface).**

So once again, you the reader have a choice, as you did in regard to the two approaches to special relativity. Einstein’s equations can be interpreted as describing a curvature of space-time, unpicturable as this may be, or as a quantum field in three-dimensional space, similar to the other quantum force fields. To the physicist, it really doesn’t make much difference. Physicists are more concerned with solving their equations than with interpreting them:

> **The important thing is to be able to make predictions about images**

> **on the astronomers photographic plates, frequencies of spectral lines, and so on, and it simply doesn't matter whether we ascribe these predictions to the physical effects of gravitational fields on the motion of planets and photons or to a curvature of space and time. (The reader should be warned that these views are heterodox and would meet with objections from many general relativists.) – S. Weinberg, (W1972, p. 147)**

You can believe that gravitational effects are caused by curvature of space-time if you want or, like Einstein, Weinberg, Wilczek (and me), you can view gravity as a force field that exists in three-dimensional space and evolves in time according to the gravitational field equations.

SUMMARY

- In QFT, space is the same three-dimensional “Euclidean” space that we intuitively believe in, and time is the same time that we intuitively believe in.
- The gravitational field is a force field like the other force fields, but with a higher spin, or helicity, of 2. Four-dimensional curvature is best left to the physicists who find it useful in their calculations.
- General relativity is compatible with QFT, at least in Schwinger’s formulation. However, unlike the equations for the EM field, the gravity field equations cannot be renormalized and calculations cannot be made at the quantum level.

CHAPTER 9
QUANTUM MECHANICS

> **A year or so ago, while Philip Candelas and I were waiting for an elevator, our conversation turned to a young theorist who had been quite promising as a graduate student and who had then dropped out of sight. I asked Phil what had interfered with the ex-student's research. Phil shook his head sadly and said, "He tried to understand quantum mechanics." — S. Weinberg (W1992, p. 84)**

Now we come to the greatest enigma of all: Quantum Mechanics (QM). The previous enigmas, special and general relativity, are understood, or at least accepted, by most physicists, but QM drives even physicists crazy. Some, like Candelas' student, decide to leave physics because of it. Another physicist — Paul Ehrenfest, a good friend of Einstein — was driven to suicide by the paradoxes of QM:

> **In recent years it has become ever more difficult for me to follow the developments in physics with understanding. After trying, ever more enervated and torn, I have finally given up in desperation. This made me completely weary of life… I tried other things but that helps only briefly. Therefore I concentrate more and more on the precise details of suicide… I have no other practical possibility than suicide… Forgive me… — P. Ehrenfest (www-history.mcs.st-and.ac.uk)**

The main reasons that QM is incomprehensible are:

- Wave-particle duality
- The uncertainty principle
- The measurement problem
- Entanglement

WAVE-PARTICLE DUALITY

The concept of wave-particle duality was introduced by Einstein in the 1905 paper that earned him the Nobel Prize. He argued that, since EM radiation is emitted in discrete units by single atoms (as Planck had shown) and is absorbed in discrete units, then surely each unit must be localized in space — like a particle. How else could it be in a position to deposit all its energy into a single atom? On the other hand, the validity of Maxwell's equations for the EM field was indisputable and the wave behavior of EM radiation was well-documented. Thus was born the idea of *wave-particle duality*.

This concept was extended to matter in 1920 by Louis de Broglie, who showed that the electron, long thought of as a particle, also exhibits wave characteristics (see Chapter 6). If the photon's particle-like behavior could not be ignored, then the electron's wave-like behavior was even less ignorable. And so Schrödinger's famous equation was not interpreted as a field equation as Schrödinger would have liked, but as an equation that gives the probability of finding a particle at a particular location.

Resolution. The wave-particle duality paradox is resolved in a very simple way. In QFT there are no particles; there are only fields. The explanation for the particle-like behavior is quantum collapse, as described in Chapters 3 and 6. Each field quantum has its own identity and acts as a unit. If a quantum is absorbed by an atom, all its energy is deposited into that atom, no matter how spread out it may be.

THE UNCERTAINTY PRINCIPLE

The probabilistic interpretation of Schrödinger's equation eventually led to the *uncertainty principle* of QM, formulated by Werner Heisenberg in 1927. This principle states that the exact position of a particle, say an electron, cannot be determined, but the uncertainty in position is related to the uncertainty in momentum by a mathematical equation. So not only do we have wave-particle duality to deal with, we have to deal with particles that might be here or might be there, but we can't say where. If the electron is really a particle, then it only stands to reason that it must be *somewhere*.

Resolution. In QFT there are no particles (stop me if you've heard this before) and hence no position – certain or uncertain. Instead there are blobs of field spread out over space. Instead of a particle that is either here *or* here *or* possibly there, we have a field that is here *and* here *and* there. Spreading out is something that only a field can do; a particle can't do it.

In fact, there is a property of fields called *Fourier's theorem* that relates the spatial spread of a field to the spread of its wavelengths. Now in QFT the wavelength of a quantum is related to its momentum, so Fourier's theorem is equivalent to the relationship between position and momentum in the Uncertainty Principle. I still remember my moment of insight in graduate school when I realized that Heisenberg's Uncertainty Principle is nothing more than Fourier's theorem.

THE MEASUREMENT PROBLEM

The "measurement problem" has been called "the most controversial problem in physics today" (informationphilosopher.com). The problem arises because Quantum Mechanics doesn't offer a picture of reality when no one is looking. Instead there are particles that are both here and there and states that are in superpositions. This was too much for Einstein and Schrödinger to swallow. Can it be that nothing really happens — that there are only probabilities — until someone looks? As Einstein lampooned, "Is the moon only there when we look at it?"

Einstein's bomb. To show how ridiculous this view is, Einstein proposed the following thought-experiment. Imagine that a keg of gunpowder is triggered by the quantum instability of some particle. The quantum mechanical equation for this situation, he said "describes a sort of blend of not-yet and already-exploded systems… but *in reality* there is just no intermediary between exploded and not-exploded" (I2007, p. 456).

Schrödinger's cat. Worried that an explosion that is only half-real might not be enough to convince people, Schrödinger proposed a thought experiment involving a cat — a cat that went on to become the most famous cat in physics history (Fig. 9-1).

Fig. 9-1. Illustration of cat both alive and dead. (wikipedia – old article)

One can even set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter

there is a tiny bit of radioactive substance, so small that perhaps in the course of the hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives *if* meanwhile no atom has decayed. The [wave-function] of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts. – E. Schrödinger (from a 1935 paper in Die Naturwissenschaften, trans. by John D. Trimmer in Proc. Am. Phil. Soc. 124, 5, 1980.)

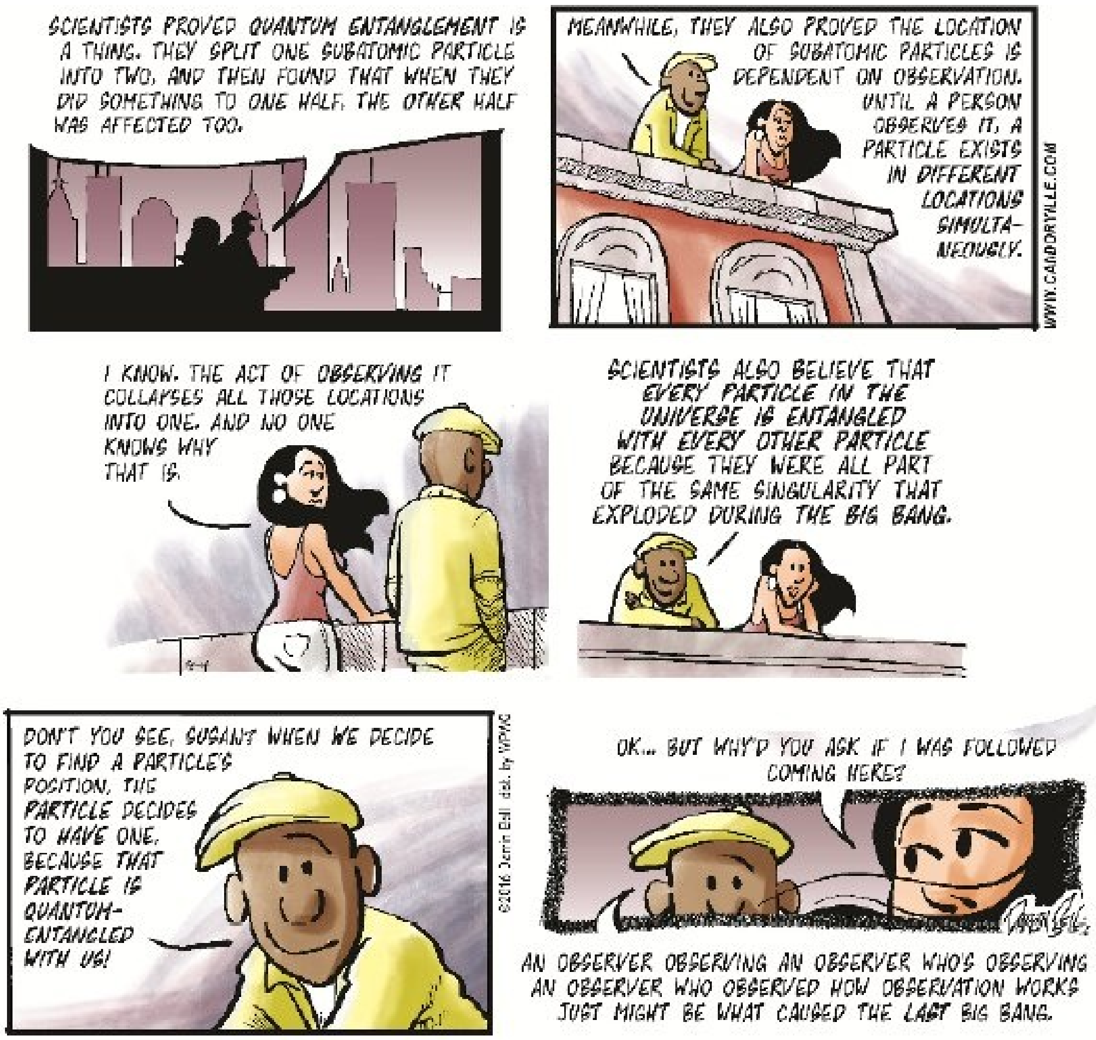
Resolution. QFT supplies a simple answer for Schrödinger's cat and Einstein's bomb. The answer, again, is quantum collapse. Quantum collapse happens with or without an observer. In the cat scenario, when a quantum is emitted from the radioactive sample, it first interacts with all other quanta that it encounters, as described by the field equations. These interactions are deterministic and reversible. This phase ends when the quantum collapses and transfers its energy to an atom in the Geiger counter. This triggers a discharge that trips the relay that releases the poison gas that kills the cat. Until then the cat is alive; after that the cat is dead. There are no superpositions of states.

ENTANGLEMENT

Quantum collapse can also occur if two quanta are created together so that their properties (spin, momentum, etc.) are interrelated. Such quanta are said to be *entangled*; if one quantum collapses or changes its state, the other must do the same and it must do it instantaneously. Experiments with entangled photons have demonstrated that when the spin of one photon changes (via interaction with a magnet), the other spin also changes, and it does it at the same time, no matter how far apart the photons are. This is what Einstein called “spooky action-at-a-distance.”

Resolution. In QM there is no way to explain this, but in QFT it is just another instance — a more elaborate instance — of quantum collapse. If one can accept that a single quantum, even if spread over miles of space, can instantaneously collapse, it is not much of a stretch to accept that two entangled quanta can do the same.

Unfortunately, the QFT resolutions to these paradoxes have not penetrated very far into public awareness. Instead, most people believe that Schrödinger’s cat *is* half-dead and half-alive, that a particle’s position *is* determined when someone looks, and that entanglement *is* “spooky”. This view is so pervasive that it can even be found on the comics page, as in this Candorville strip by Darrin Bell, published on Sept. 4, 2016:


SCIENTISTS PROVED QUANTUM ENTANGLEMENT IS A THING. THEY SPLIT ONE SUBATOMIC PARTICLE INTO TWO, AND THEN FOUND THAT WHEN THEY DID SOMETHING TO ONE HALF, THE OTHER HALF WAS AFFECTED TOO.
MEANWHILE, THEY ALSO PROVED THE LOCATION OF SUBATOMIC PARTICLES IS DEPENDENT ON OBSERVATION. UNTIL A PERSON OBSERVES IT, A PARTICLE EXISTS IN DIFFERENT LOCATIONS SIMULTANEOUSLY.
WWW.CANDORVILLE.COM
I KNOW. THE ACT OF OBSERVING IT COLLAPSES ALL THOSE LOCATIONS INTO ONE. AND NO ONE KNOWS WHY THAT IS.
SCIENTISTS ALSO BELIEVE THAT EVERY PARTICLE IN THE UNIVERSE IS ENTANGLED WITH EVERY OTHER PARTICLE BECAUSE THEY WERE ALL PART OF THE SAME SINGULARITY THAT EXPLODED DURING THE BIG BANG.
DON'T YOU SEE, SUSAN? WHEN WE DECIDE TO FIND A PARTICLE'S POSITION, THE PARTICLE DECIDES TO HAVE ONE, BECAUSE THAT PARTICLE IS QUANTUM-ENTANGLED WITH US!
©2016 Darrin Bell dist. by WPWG
OK... BUT WHY'D YOU ASK IF I WAS FOLLOWED COMING HERE?
AN OBSERVER OBSERVING AN OBSERVER WHO'S OBSERVING AN OBSERVER WHO OBSERVED HOW OBSERVATION WORKS JUST MIGHT BE WHAT CAUSED THE LAST BIG BANG.

SUMMARY

In QFT the paradoxes of QM have simple answers:

- There is no wave-particle duality because there are no particles. The particle-like behavior is explained by the fact that field quanta live and die as a unit. This phenomenon is called quantum collapse.
- The Uncertainty Principle is simply a statement that fields cannot be localized; they spread out.
- The measurement problem is solved by quantum collapse. There is no role of the observer. Quantum collapse happens whether or not someone is looking.
- Entanglement is a more elaborate example of quantum collapse that occurs when two quanta are created together.

Is that all there is to it? Did I give too little space to discussing these "profound" paradoxes? Well, that's really all there is to it. In QFT everything is fields. They spread out, they collapse, and they do all this without requiring an observer. When I hear people complaining about the weirdness and inaccessibility of modern physics, I want to ask, "What part of Quantum Field Theory don't you understand?"

CHAPTER 10

THE TRIUMPH OF QUANTUM FIELD THEORY

In this chapter we present an overview of Quantum Field Theory. We will describe its structure and some of its many successes. Also, out of fairness, we will describe the gaps in the theory. But first we will start with the three pillars on which this amazing edifice rests, as presented by Julian Schwinger in his lectures at Harvard University in 1956-59.

THE PILLARS

The field principle. The first pillar is the assumption that nature is made of fields. A field is a set of physical properties that exist at every point of space. However, the concept of a field as a property of space does not come easily. It eluded the great Newton, even though he couldn't accept action-at-a-distance. It wasn't until 1845 that Faraday, inspired by patterns of iron filings, conceived the idea of fields, and it took another 50 years before the concept was accepted without invoking an imaginary ether. QFT comprises seven fields — five force fields and two matter fields.[1] The force fields include gravity, electromagnetic forces, strong and weak nuclear forces, and the recently-discovered Higgs field. The matter fields include lepton and baryon fields. Two of these fields — strong and baryon — are *effective* fields that are made of more basic but "invisible" fields called quarks and gluons. The use of colors is my attempt to make the field picture more palatable.

The relativity principle. The second pillar is the assumption that the field equations are the same for all uniformly-moving observers. This is Einstein's Principle of Relativity, famously enunciated in 1905. QFT is the only theory that successfully combines the relativity and quantum principles.

The quantum principle (discretization). The quantum principle was introduced in 1900 by Max Planck, who showed that EM radiation emitted by hot objects consists of discrete chunks that he called quanta (Chapter 3). Discretization was demonstrated experimentally in 1922 by Otto Stern and Walther Gerlach. Their classic experiment showed that the angular momentum (or spin) of the electron can have only two values — nothing in between (Fig. 10-1). In Schwinger's QFT all physical properties are treated as discrete (S2001). Even field strengths, whose values are continuous, are treated mathematically as the limit of increasingly finer discrete values. It is discretization that leads to the use of Hilbert algebra as the language of QFT.

Fig. 10-1. The Stern-Gerlach experiment. A beam of atoms is deflected when it enters a non-uniform magnetic field. Classically one would expect that the atom's magnetism (which arises from the spin of its electrons) could have any magnitude and that the beam would be deflected into a continuous band. Instead, the beam separates into two distinct parts as shown, corresponding to discrete spin values of +½ and -½ in Planck units. (www.upscale.utoronto.ca)

Occam's Razor. I'm tempted to add another principle, but it's really more of a wish than a rule. I'm referring to Occam's razor, which states in essence that all other things being equal, the simplest explanation is the best. Einstein put it somewhat differently: "A physical theory should be as simple as possible, but no simpler." The last phrase is important because, as Schwinger said, "nature does not always select what we, in our ignorance, would judge to be the most symmetrical and harmonious possibility". If the theory were as simple as possible, there would be just one field (or perhaps none!), and the world would be very uninteresting — not to mention uninhabitable. I think it can be said that the equations of QFT are indeed about as simple as possible, but no simpler.

> **The move from a particle description to a field description will be especially fruitful if the fields obey simple equations, so that we can calculate the future values of fields from the values they have now… Maxwell's theory of electromagnetism, general relativity, and quantum chromodynamics all have this property. Evidently, Nature has taken the opportunity to keep things relatively simple by using fields. – F. Wilczek (W2008, p. 86)**

THE EDIFICE

On these pillars rests the most successful theory ever constructed. Except for a few gaps, this theory explains everything from the tiniest atomic nucleus to the most remote star. Not only that, but most of the explanations emerge from the theory as easily and naturally as raindrops falling from the clouds, or better yet, like presents appearing under the Christmas tree. Following are some of the more attractive and accessible of these presents. Some have already been mentioned and some will be pleasant surprises.

Quanta. Quantum fields exist in three different forms: quanta, self-fields, and vacuum fields. A quantum is a separate, indivisible chunk of field that lives a life and dies a death of its own. For example, the photon is a quantum of the EM field, and protons and neutrons are quanta of the baryon field. Quanta are sometimes called *excitations* in a field, but that term doesn't do them justice. Excitations can have any magnitude and can diminish as they travel and slowly die away like water waves, whereas quanta are indivisible and act as a unit. They may be free and travel through space on their own, or they may be bound, as an electron in an atom, but each quantum keeps its own identity. If it is absorbed or changes its spin state, it does so as a unit. Because of this all-or-nothing behavior, quanta act like particles, making it difficult for many physicists to give up their belief in particles.

Self-fields. Self-fields do not have a life of their own; they are created by a source and are permanently attached to that source (see Fig. 3-1). Examples of self-fields are the gravitational field of the earth, the electric field around an electron or proton, and the strong and weak fields around a nucleon.

The vacuum field. Finally, there is the vacuum field. The equations of quantum field theory do not permit the field strength ever to be zero. Even in regions where there are no quanta or self-fields, there is a background field called the vacuum field. The vacuum field is especially important in the case of the Higgs mechanism (see below).

Hilbert algebra. Quantum fields are not described by simple numbers. They are described by vectors in what mathematicians call *Hilbert space* and their dynamics are described by *operators* that obey partial differential equations. Thus, while classical field intensity is described by a simple number in QFT we talk about the expectation value of field intensity. However, since my aim is to avoid mathematics, I will not go into this any further. You do not need to understand Hilbert algebra, or even to know the field equations, to grasp the basic concepts of QFT. Just remember that Hilbert space is not real; it is a mathematical tool and is not to be confused with the physical fields that exist in real three-dimensional space.

Mass. In classical physics mass is a measure of inertia (see Chap. 2), but in QFT mass is a number that appears in the field equations. The effect of the mass term is to slow down the speed at which a field evolves and propagates, so mass plays the same inertial role in QFT that it does in classical physics. But this is not all that mass does. This same term also causes the fields to oscillate, and the greater the mass, the higher the frequency.[2]

Energy. In classical physics, energy means the ability to do work and work is defined as force exerted over a distance. This definition, however, doesn't provide much of a picture, so in classical physics energy is a rather abstract concept. In QFT, on the other hand, the energy of a quantum is represented by oscillations in its field intensity: the more energy, the faster the oscillations. In fact Planck's famous relationship between energy and oscillation frequency (see Chap. 3) follows directly from the equations of QFT. In our color analogy, we might say that the oscillations cause the color to "shimmer"; the faster the shimmer, the greater the energy of the field.

$e = mc^2$. I know, I promised there would be no equations and, except for a few footnotes, I've kept my promise. But I think you will forgive me for making an exception for the world's most famous equation — the only equation to have its biography written (B2000). And the thing is this: $e = mc^2$ pops right out of QFT. Einstein had to work hard to derive it; it was published in a separate paper that followed his first breakthrough paper on relativity theory in 1905, but in QFT this equation follows as a simple consequence of the two previous results. Since both mass and energy are represented by oscillations in the field intensity, it doesn't take an Einstein to see that there is a relationship between the two. In fact, any schoolboy can combine the two equations and find (big drum roll, please) $e = mc^2$.[3] Not only does this equation tumble right out of QFT, its meaning is seen physically in the oscillations of the fields. (This simple derivation only occurred to me as I was writing this book.)

Field components. Each of the seven fields in QFT is actually a group of interrelated "sub-fields" that share a set of equations. For example, what we call the EM field is a combination of electric and magnetic fields that are described by Maxwell's equations. Another example is the lepton field, which consists of four component fields to accommodate the two possible spins and charges

Spin (or helicity). In QFT there are no particles, so there is nothing to spin on its axis and the concept of *spin* is not easy to picture. Because of this, the term *helicity* is sometimes used in place of spin. Helicity is a mathematical concept related to the number of field components and how they change when viewed from different angles. Helicity gives rise to angular momentum, just as does a spinning particle.

Exclusion principle. The *exclusion principle* states that two field quanta with half-integer spin (called *fermions* after Enrico Fermi) cannot be in the same quantum state. This explains why matter quanta (electrons, protons, etc.) are seen only as separate entities. It also explains why each electron in an atom must be in a different quantum state (as described in Chapter 6) and why there is a limit to the number of neutrons in an atom (see "A delicate balance" in Chapter 4).

When the exclusion principle was introduced by Wolfgang Pauli in 1925, it was only a guess — an empirical postulate. It is only in QFT that this important principle has a theoretical basis as a consequence of the *spin-statistics theorem*.

> **In my original paper I stressed the circumstance that I was unable to give a logical reason for the exclusion principle or to deduce it from general assumptions… If we search for a theoretical explanation of this law, we must pass to the discussion of relativistic wave mechanics. – W. Pauli (Nobel lecture, 1945)**

Classical limit. The exclusion principle does not apply to force field quanta (called *bosons* after Satyendra Bose), which have integer spin. They can overlap and build up. Even though each quantum acts individually, if there are many present the effect is the same as a classical field.

Quantum collapse. If fields evolved only as described by the field equations, nothing of significance would happen because the equations do not describe the transfer of energy or momentum. For example, they don't describe how a photon transfers its energy to a photoreceptor in your eye. I call this process *quantum collapse* because after a quantum transfers its energy it can't continue to exist; it must disappear from all space. Like Schrödinger's cat, it can't be in a state of half dead and half alive. Even though we have no theory to describe quantum collapse, we know that it happens; indeed, it is an essential part of QFT.[4] As Art Hobson wrote, referring to the two-slit experiment:

> **The entire spread-out field… must deposit its quantum of energy all at once, in a single instant, because the field cannot carry some fraction of one quantum – it must always contain either exactly one or exactly zero quanta of energy. When the field deposits its quantum of energy on the viewing screen, the entire spread-out field must instantaneously lose this much energy. — A. Hobson (H2007, p. 303)**

Quantum collapse is similar to collapse of the wave-function in QM, but the concepts are very different. In QM the wave-function describes the probability that a particle is in a given place or that the system is in a given state, and when a measurement is made the probabilities "collapse" into a certainty. In QFT the collapse is a *physical process* that actually happens. It may not be what we expected, but neither did we expect that the earth is round or that matter is made of atoms. Just as we learned to live with those concepts, so we can learn to live with quantum collapse.

Not only is quantum collapse a necessary part of QFT, it provides solutions to two of the most vexing problems in physics today: the measurement problem and entanglement, as described in Chapter 9.

Quarks and gluons. Quarks and gluons are the basic fields that constitute the strong and baryon fields, but they do not exist in free form. This is known as the Principle of Confinement. Indeed, we would not have known about them if not for the large number of *hadrons* (quanta of the strong and baryon fields) that were discovered in the latter half of the 20th century – the so-called "subatomic zoo". It was Yuval Ne'eman, an Israeli physicist, and Murray Gell-Mann who saw a pattern in the "zoo" that Gell-Mann dubbed "The Eightfold Way". This pattern led Gell-Mann to predict a new particle (the omega-minus) that was observed in 1963 with the predicted properties. Then in 1964 Gell-Mann and George Zweig showed that this pattern would result if hadrons were made of more basic fields that Gell-Mann (quite the wordsmith) called *quarks*, a word taken from Ulysses by James Joyce.[5] Later *gluons* (again, Gell-Mann's term) were introduced to hold the quarks together. The field theory that describes quarks and gluons was given the name (by guess whom) *quantum chromodynamics* (QCD), because arbitrary colors are used to describe certain properties of the quarks. (These colors are not to be confused with the equally arbitrary colors I have used to help visualize fields.) While

QCD has its own name, it still is part of Quantum Field Theory.

<u>"One of the greatest scientific achievements of all time"</u>. In a recent calculation it was shown that the field equations for quarks and gluons account for the mass and spin of all the hadrons in the sub-atomic zoo.[6] The basic parameters of the quark and gluon fields were first determined from three of the known masses. The field equations were then used to calculate stable and quasi-stable excitations of the quark and gluon fields. Fourteen such excitations were found with mass and spin in close agreement with the fourteen known hadrons, ranging from the proton and neutron to the exotic charmonium. Equally important is that there were no excitations corresponding to the quarks and gluons themselves, thereby providing a theoretical basis for the Principle of Confinement. As Frank Wilczek wrote:

> **Through difficult calculations of merciless precision that call upon the full power of modern computer technology, [we have] shown that unbendable equations… account convincingly and in quantitative detail for the existence of protons and neutrons, and for their properties… I believe this is one of the greatest scientific achievements of all time. — F. Wilczek (W2008, p. 122-127)**

An achievement, I might add, that very few people are aware of.

<u>The Higgs field</u>. The Higgs field is the newest field in QFT; its quantum (called the Higgs boson) was detected in 2012. It interacts with all other fields and, because it has a large vacuum expectation value, it generates the effective mass of these fields. It was the final ingredient that made electroweak unification possible (see Chapter 5).

<u>Dark Matter</u>. As early as 1933, astronomical observations showed that there is more matter in the universe than was previously thought. In fact, astronomers now believe that ordinary matter constitutes only about 5% of the total mass of the universe, while something called *dark matter* makes up 27% of the total mass. (The rest is made of a related substance called dark energy, which we won't go into.) This conclusion is based on its apparent gravitational influence on other objects.[7] Dark matter is found primarily around galaxies and is believed to be a million times less dense than normal matter. It is not "seen" because it doesn't emit or absorb light — that is, it doesn't interact (or interacts very feebly) with the EM field.

QFT offers two possible explanations for dark matter. One possibility is a field suggested by Steven Weinberg and Frank Wilczek, with quanta called axions. Attempts have been made to detect axions with an apparatus called DAMA (for DArk MAtter.) that is buried under the Gan Sasso mountain in Italy. According to the DAMA scientific team, a signal was detected that is consistent with dark matter, but other scientists have treated this finding with skepticism and new experiments are underway. The other possibility is a field suggested by an approach to QFT called *supersymmetry*. Either of these hypothesized fields would be sufficiently stable and interact feebly enough with normal matter to qualify. To go further into dark matter, axions, and supersymmetry is beyond the scope of this book (and its author), but I would like to quote the delightful Frank Wilczek as to the origin of the name axion:

> **I'm very fond of axions, in part because I got to name them. I used that opportunity to fulfill a dream of my youth. I'd noticed that there was a brand of laundry detergent called "Axion", which sounded to me like the name of a particle. So when theory produced a hypothetical particle that *cleaned up* a problem with an *axial* current, I sensed a cosmic convergence. The problem was to get it past the notoriously conservative editors of *Physical Review Letters*. I told them about the axial current, but not the detergent. It worked. – F. Wilczek (W2008, p. 203)**

THE GAPS

Now we will take a look at what is not covered by QFT — the gaps in the theory.

Renormalization. The most obvious gap is that QFT does not describe the interaction of a quantum with its self-field—an interaction that Richard Feynman called a "silly" idea (Chapter 6). The problem is that according to QFT the result of this interaction turns out to be infinite, which is obviously not correct. The problem was solved for EM fields by the simple expedient of replacing the infinite quantities by their experimentally-determined values. However, the fact that the infinities occur in the first place shows that something is going on at the single quantum level that is not described by the QFT equations. As we saw in Chapter 8, this problem is more serious for the gravity field because its equations do not allow for the renormalization solution.

Quantum collapse. While quantum collapse is an essential part of QFT, it is not described by the field equations. Besides this lack of theory, there are several aspects of it that bother many physicists. One is that the collapse is instantaneous and occurs at the same time at widely separated points.[8] This is especially bothersome in the case of two entangled photons. Physicists call such a process *non-local* because it involves communication faster than the speed of light. Now it is true that the field equations include a number *c* that limits the speed of propagation, but quantum collapse is not described by these equations, so there is no reason it can't occur. In fact, quantum collapse is necessary if quanta are to act as indivisible units. Since it does not lead to any paradoxes or inconsistencies, there is no reason not to accept it. In any event, non-locality has been experimentally documented. Even those who believe in particles as the ultimate reality acknowledge that something collapses.

The other problem is that, so far as we know, quantum collapse is random. QFT does not explain when or how it happens. All we know is that the probability of collapse is related to the field intensity at a given point. The idea of randomness was troubling to Einstein:

> **I find the idea quite intolerable that an electron exposed to radiation should choose, of its *own free will,* not only its moment to jump off, but also its direction. In that case, I would rather be a**

cobbler, or even an employee in a gaming house, than a physicist. – A. Einstein (letter to Max Born, 1924)

However, 25 years later Einstein softened his stand. In a 1950 speech to the International Congress of Surgeons, after describing the "overwhelming evidence" for giving up causality, he concluded by saying:

Will this credo survive forever? It seems to me a smile is the best answer. – A. Einstein (Physics Today, June 2005, p. 47-48)

While the problem of randomness is not solved by QFT, at least it has been pinned down to a specific event. It is no longer a vague phenomenon related to the role of the observer, as in QM; it is a physical event that happens with or without an observer. Maybe someday we will have a theory to describe it, but even if it is truly random there is nothing inherently contradictory about that. It may not be what we expected, but, like Einstein, we can always smile.

A speculation. Quantum collapse and renormalization both involve something that happens at the single quantum level – something not described by QFT. Renormalization is needed because QFT doesn't explain how a quantum interacts with its self-field. Quantum collapse is a mystery because QFT doesn't describe when or how a quantum transfers energy or momentum. It is possible that these two problems are related, and that if one gap is filled, the other will also be filled.

The next two gaps are not unique to QFT; they are present in every theory.

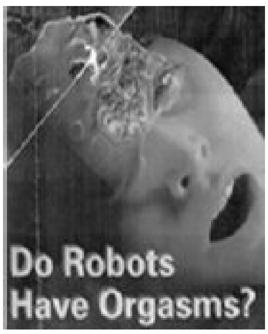
Whys and Wherefores. QFT does not explain why the numbers that appear in the equations have the values they do. The most famous example is the so-called *fine structure constant* that describes the interaction between matter fields and the EM field. This constant was once thought to have a value of 1/137, and this, as you might imagine, led to some numerological attempts to explain why nature had chosen this particular integer — and such an unusual one at that. Sir Arthur Eddington claimed that the number could be obtained by "pure deduction", but these attempts were abandoned when more precise measurements showed that the actual value is 1/137.04.

Many physicists still wonder why the masses and coupling constants are what they are, and there are attempts to find explanations that are more sophisticated than the 1/137 saga, including something called *superstring theory*. There are also less sophisticated attempts, such as the so-called *anthropic principle*, which states that if the values were different from what they are, the human race could not exist. QFT does not supply answers, nor does any other theory. QFT does an amazing job of explaining the world we live in, but why the constants are what they are is, in my opinion, a teleological, if not theological, question.

Consciousness. Finally we come to the grand-daddy of mysteries. How dare physicists talk about "theories of everything" when they can't explain what goes on behind their very noses! But please understand, by consciousness I don't mean simple information processing, such as can be done by any computer. I mean the sense of awareness, the sensations, the feelings that human and other minds experience every day – from the color red to the beauty of a Mozart sonata or the pain of a toothache. Such sensations are known as *qualia*. Most physicists don't want to be bothered with the question, and it is left to philosophers like Charlie Chaplin to worry about it:

> **Billions of years it's taken to evolve human consciousness… The miracle of all existence… More important than anything in the whole universe. What can the stars do? Nothing but sit on their axis! And the sun, shooting flames 280,000 miles high. So what? Wasting all its natural resources. Can the sun think? Is it conscious? — C. Chaplin (film "Limelight")**

I see consciousness as a more pressing problem than the question of why the field constants have the values they do. My concern began when I was thinking about pain. Pain surely cannot be explained by fields or particles or relativity, or even quantum field theory, I thought. Nor can pleasurable sensations, like the enjoyment of music or the intense pleasure of an orgasm.



Then one day we had a visitor - a young computer hot shot. As we were sitting down for dinner I asked him "Do you think a computer can ever experience a sexual orgasm?" Well this young fellow began to tell me how you could create an orgasm by putting 0's and 1's into the right memory banks. Of course this was nonsense, so I told him he had flunked the test and couldn't have any dinner. I didn't really; we fed him, but it gave me the idea to write a book about consciousness with the title "Do Robots Have Orgasms?" (The illustration shows the cover designed by my nephew Roger L. Brooks.) I eventually abandoned the book because I figured that a book that boiled down to one word ("no") wouldn't sell. However, I had already written or sketched chapters that I called "dead ends" — three explanations that have been proposed for consciousness: Artificial Intelligence, Religion, and Quantum Mechanics. As I was working on the chapter about Quantum Mechanics, I realized that all the quantum mechanical explanations ignored QFT. In fact, I became aware that QFT is ignored almost everywhere — as if it never existed. And that's why I wrote the book that I wrote.

Is there any prospect that the consciousness mystery will ever be solved? Ambrose Bierce didn't think so. To quote from his Devil's Dictionary:

> **Mind, n. A mysterious form of matter secreted by the brain. Its chief activity consists in the endeavor to ascertain its own nature, the futility of the attempt being due to the fact that it has nothing but itself to know itself with. – Ambrose Bierce (B1958, p. 87)**

SUMMARY

Among physicists there are many different approaches to understanding reality. Some physicists believe that reality consists of particles, despite the many inconsistencies and absurdities, not to mention questions like what the particles are made of. Some believe in wave-particle duality, which is neither fish nor fowl (see Norsen quote on p. 4) . And others, like Steven Hawking (see quote on p. 4) don't worry about reality. As Steven Weinberg said,

> **It is truly surprising how little difference all this makes. Most physicists use quantum mechanics every day in their working lives without needing to worry about the fundamental problem of its interpretation. Being sensible people with very little time to follow up all the ideas and data in their own specialties and not having to worry about this fundamental problem, they do not worry about it. — S. Weinberg, (W1992, p. 84)**

But for those who believe there is a reality and who want to understand it, the choice was described this way by Robert Oerter:

> **Wave or particle? The answer: Both, and neither. You could think of the electron or the photon as a particle, but only if you were willing to let particles behave in the bizarre way described by Feynman: appearing again, interfering with each other and cancelling out. You could also think of it as a field, or wave, but you had to remember that the detector always registers one electron, or none — never half an electron, no matter how much the field has been split up or spread out. In the end, is the field just a calculational tool to tell you where the particle will be, or are the particles just calculational tools to tell you what the field values are? Take your pick. — R. Oerter (O2006, Chap. 6: "Feynman's Particles, Schwinger's Fields", p. 128)**

But before you take your pick, let us take a look at some of the things QFT has accomplished:

- QFT explains why the detector always register one electron or none (quanta are indivisible).
- QFT provides a simple derivation of $e = mc^2$ and gives it a meaning (both are oscillations in a field).

- QFT explains the Pauli Exclusion Principle (the spin-statistics theorem).
- QFT explains why matter quanta act like particles (the Exclusion Principle).
- QFT explains why the number of neutrons in a nucleus is limited (again, the Exclusion Principle).
- QFT explains why force fields become classical fields in the limit of many quanta (the Exclusion Principle doesn't apply).
- QFT explains the Higgs mechanism (it's another field).
- QFT offers two possible explanations for dark matter (two other fields).
- QFT explains the *subatomic zoo* ("one of the greatest scientific achievements of all time").
- QFT explains the paradoxes of special relativity (a natural consequence of the way fields behave).
- In QFT time is different from space (in accord with our natural perception).
- QFT is compatible with general relativity (although there are calculational difficulties).
- In QFT (and in general relativity) gravity is a force field, not curvature of space-time.
- QFT (and general relativity) explains gravity waves as oscillations in this field.
- QFT explains how these waves were detected (gravity-induced contraction).
- QFT explains wave-particle duality (there are no particles, there are only fields).
- QFT explains the Uncertainty Principle (fields spread out in space)
- QFT offers a solution to the measurement problem (quantum collapse).
- QFT offers an explanation of Einstein's "spooky action at a distance" (entangled quanta collapse).

Well, that's quite a list of accomplishments! With all that, you must surely wonder why QFT hasn't been accepted, if not embraced, by the physics community and the public. Well, there is a downside. To reap those benefits, we must accept that:

- Quantum fields are described mathematically by vectors in Hilbert space, not by simple numbers.
- QFT doesn't tell us how a quantum interacts with its self-field.
- QFT doesn't tell us why or when quantum collapse occurs.
- Quantum collapse is instantaneous (i.e., non-local).

And so, dear reader, I hope that, like Frank Wilczek, Steven Weinberg, Sean Caroll, Art Hobson, Julian Schwinger (and me), you will choose QFT: the only theory that offers a picture of reality that is understandable and makes sense. And perhaps one day the physics community will finally abandon the QM ship made of particles floating on a sea of paradox for smoother sailing on the seas of quantum fields.

BIBLIOGRAPHY

A1982. Asimov's Biographical Encyclopedia of Science and Technology by Isaac Asimov, Doubleday & Co., Garden City, NY, 2nd rev. edition. If you have only one book on science in your library, this should be it.

A2004. Math and the Mona Lisa by Bülent Atalay, Smithsonian Books, Washington.

B1958. The Devil's Dictionary by Ambrose Bierce, Dover, NY. Written between 1881 and 1906 and originally called The Cynic's Word Book. "Among the series of epigrams called The Devil's Dictionary are some of the most gorgeous witticisms in the English language" — H. L. Mencken

B1966. The World of the Atom, ed. by Henry A. Boorse and Lloyd Motz, Basic Books. An anthology of original articles and commentary about atomic theory from 65 BC to 1965.

B1983. The Birth of Particle Physics, ed. by Laurie M. Brown and Lillian Hoddeson, Cambridge University Press.

B2000. $e = mc^2$: A biography of the world's most famous equation by David Bodanis, Pan Books.

B2001. The Foundations of Quantum Mechanics by John S. Bell, World Scientific, Singapore. A collection of Bell's articles about QM plus two articles on relativity: "How to teach special relativity" and "In memory of George Francis FitzGerald."

B2003. A Short History of Nearly Everything by Bill Bryson, Broadway Books, NY. A highly readable book about science in general.

C2012. The Particle at the End of the Universe by Sean Carroll, Penguin Group

D1958. Quantum Mechanics by Paul A. M. Dirac, Oxford University Press.

D2004. Discover magazine, vol 25, No. 9, Sept. 2004 (Special Einstein Issue)

E1923. The Principle of Relativity by Albert Einstein et al., transl. by W. Perrett and G. B. Jeffery, Dover Publications. A collection of original papers about the special and general theories.

E1961. Relativity: The special and general theory by Albert Einstein, Three Rivers Press, NY. Einstein's own "popular" presentation, recommended for those who can tolerate the 37 equations.

E1956. Out of My Later Years by Albert Einstein, Citadel Press, NY. A collection of essays on science, world politics, etc.

E1979. Albert Einstein: The Human Side, ed. by H. Dukas and B. Hoffan, Princeton University Press. A very readable collection of excerpts from Einstein's letters about many subjects from Zionism to intermarriage.

F1985. Surely You're Joking, Mr. Feynman (Adventures of a curious character) by Richard Feynman, W.W. Norton & Co. Delightful and earthy anecdotes of his personal and professional life, as told to Ralph Leighton. Note the double meaning of the word "curious."

F1985a. QED: The strange theory of light and matter by Richard Feynman, Princeton University Press. Feynman's attempt to explain Quantum Mechanics to the "masses".

F1988. What Do You Care What Other People Think? (Further adventures of a curious character) by Richard Feynman, W.W. Norton & Co.

G1992. Genius: The life and science of Richard Feynman by James Gleick, Vintage Books, NY. The definitive biography of Feynman.

G2004. The Fabric of the Cosmos: Space, time, and the texture of reality by Brian Greene, Vintage Books, NY.

H2007. Physics: Concepts and Connections by Art Hobson, 4th ed., Pearson Prentice Hall, NJ. A remarkable college textbook for non-science majors that embodies the QFT approach.

H2015. Tales of the Quantum: Understanding physics' most fundamental theory" by Art Hobson, Oxford University Press, to be published

I2002. The Man Who Knew Too Much: The strange and inventive life of Robert Hooke, by Stephen Inwood, MacMillan

I2007. Einstein: His Life and Universe by Walter Isaacson, Simon & Schuster, NY. Thanks partly to 130 personal letters made public in 2006, this is the most comprehensive and definitive biography of a most human genius.

J2000. "The Jaguar and the Fox" by George Johnson, The Atlantic, July 2000 (www.theatlantic.com)

L1915. The Theory of Electrons by H. A. Lorentz, Dover, Mineola, NY, 2nd ed. (1952). Based on a lecture series delivered at Columbia University in 1906, with 109 pages of notes added in the second edition.

M1994. The beat of a different drum: The life and science of Richard Feynman by J. Mehra, Clarendon Press Oxford

M2000. Climbing the Mountain: The scientific biography of Julian Schwinger by Jagdish Mehra & Kimball A. Milton, Oxford University Press. A monumental depiction of the life and work of the "gentle genius" of physics. The mountain in the title refers to Quantum Field Theory.

M2003. Some Time With Feynman by Leonard Mlodinow, Penguin Books. (American title: Feynman's Rainbow). A young physics professor in his

first year at Caltech meets Feynman in his last year.
M2005. It's About Time: Understanding Einstein's Relativity by N. David Mermin, McGraw-Hill. A witty and readable presentation of relativity that acknowledges the bottom-up method. Note the double meaning of the title.
N1979. Opticks by Sir Isaac Newton. Dover Publications, New York (2nd edition).
N1996. Julian Schwinger: The physicist, the teacher, and the man, ed. by Y. Jack Ng, World Scientific, Singapore. A collection of fifteen talks given at symposia to commemorate Schwinger following his death in 1994.
O2006. The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics by Robert Oerter, Plume (a member of Penguin Group), London
P1982. 'Subtle is the Lord…': The science and the life of Albert Einstein by Abraham Pais, Clarendon Press, Oxford. This biography includes the best description of Einstein's theories that I know of. The title comes from a comment by Einstein that concludes with "but malicious He is not."
P1994. Shadows of the Mind by Roger Penrose, Oxford University Press. A thoughtful and provocative book with speculative ideas about quantum collapse and consciousness.
P1997. The Large, the Small and the Human Mind by Roger Penrose et al., Cambridge University Press. Penrose's 1995 Tanner lectures, with discussion by Abner Shimony, Nancy Cartwright and Stephen Hawking.
R2003. Everything's Relative by Tony Rothman, John Wiley & Sons, Hoboken. Subtitled "And other fables from science and technology."
S1957. "A Theory of the Fundamental Interactions" by Julian Schwinger, Annals of Physics: 2, 407-434, 1957. Schwinger's magnum opus on QFT that is often ignored because of a few near-misses, despite the many hits.
S1958. Selected Papers on Quantum Electrodynamics ed. by Julian Schwinger, Dover, NY. A collection of 34 papers in their original languages, from 1927 to 1953.
S1970. Particles, Sources, and Fields, Vol 1 by Julian Schwinger, Perseus Books, Reading, Mass. Schwinger's exposition of his source theory approach to QFT, the success of which "convinced me, if no one else, of the great advantages in mathematical simplicity and conceptual clarity that its use bestowed."
S1986. Einstein's Legacy by Julian Schwinger, Dover Publications, Mineola, NY. A spinoff from an unsuccessful BBC television series, it is a careful,

clear, but somewhat mathematical presentation of Einstein's contributions at a popular level.
S1994. QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga by Silvan S. Schweber, Princeton University Press
S2001 Quantum Mechanics: Symbolism of Atomic Measurements by Julian Schwinger, ed. Byy Berthold-Georg Englert, Springer Verlag
S2004. Special issue: Beyond Einstein, Scientific American, vol 291, no. 3, September 2004.
W1910. A History of the Theories of Ether and Electricity, vol. 1 by Edmund Whittaker, Thomas Nelson, London.
W1972. Gravitation and Cosmology: Principles and applications of the general theory of relativity by Steven Weinberg, John Wiley and Sons
W1992. Dreams of a Final Theory by Steven Weinberg, Pantheon Books
W2001. Facing Up: Science and its cultural adversaries by Steven Weinberg, Harvard Univ. Press. A collection of "popular" essays by a co-winner of the 1979 Nobel Prize
W2008. The Lightness of Being: Mass, ether, and the unification of forces by Frank Wilczek, Basic Books, NY. A brilliant presentation of Quantum Chromodynamics and speculations about the future.
W2015. A Beautiful Question: Finding Nature's Deep Design by Frank Wilczek, Penguin Press, NY
Z2003. Quantum Field Theory in a Nutshell by A. Zee, Princeton Univ. Press.

Chapter 1

[1] It should be noted, however, that Einstein was only human and made his share of errors. A little-known but glaring example, never admitted by Einstein, occurred in 1915 when he threw out valid experimental data in order to obtain agreement with a theory that was later found to be incorrect (P1982, p. 248).

[2] The theory known as *special relativity* was originally called simply *relativity*. The word special was added later to distinguish it from Einstein's theory of gravity, which he called general relativity. In this book the word *relativity* by itself will refer to the special theory.

Chapter 2

[1] The actual period of revolution is 27.3 days, but because of the earth's revolution around the sun, the apparent period is longer.

[2] For those who took high school geometry, the equation is $A = 4\pi r^2$.

[3] His mistake might have been to use the incorrect value of the earth's diameter obtained by Poseidonius, rather than Eratosthenes' larger correct value. This was the same error that led Columbus into undertaking a voyage to China that he surely wouldn't have attempted if he knew the true distance.

[4] Except in this book, from which the mathematics is indeed extricated (except for occasional footnotes) so that you can understand and appreciate the *concepts* of QFT.

[5] As William of Occam (1285-1349) said, "No more things should be presumed to exist than are absolutely necessary."

[6] When Einstein eventually obtained a position at the University of Zürich, a faculty report stated: "Dr. Einstein is an Israelite and… to the Israelites among scholars are ascribed (in numerous cases not entirely without cause) all kinds of unpleasant peculiarities of character, such as intrusiveness, impudence, and a shopkeeper's mentality." (P1982, p. 185)

[7] When Einstein's great battle was over, he acknowledged Grossmann's help as follows: "Finally, grateful thoughts go at this place to my friend the mathematician Grossmann, who by his help not only saved me the study of the relevant mathematical literature, but also supported me in the search for the field equations of gravitation." (P1982, p. 213)

[8] Einstein had a disputatious streak. In 1908 he wrote an angry letter to Johannes Stark, who had published a paper that contained $e = mc^2$ without, he

felt, proper attribution. It was only after Stark responded apologetically that Einstein's gentle nature returned and he replied: "Even if I had not already regretted before receipt of your letter that I had followed the dictates of petty impulse in giving vent to that utterance about priority, your detailed letter really showed me that my over-sensitivity was badly out of place. People who have been privileged to contribute something to the advancement of science should not let such things becloud their joy over the fruits of common endeavor." (S1986, p. 70)

[9] In 1849, Armand Fizeau obtained the first accurate value for *c* by sending a beam of light from one hilltop to another. Galileo had tried this method many years earlier, but Fizeau succeeded by using mirrors and rotating disks.

[10] The equatorial bulge was created when the earth was formed. Because of the centrifugal force at the equator, the coalescing gases weren't pulled inward as strongly as at the poles, leaving the earth with a bit of a "beer belly."

[11] A similar thing happened to Albert Michelson, whose experiment that became the cornerstone of special relativity was ignored by the Nobel committee (see Chapter 7).

[12] except in quantum collapse, which is described in Chapter 10.

Chapter 3

[1] Bierce also defined electricity as "the power that causes all natural phenomena not known to be caused by something else."

[2] You can see this notebook today at Faraday's laboratory in London, now a museum.

[3] Maxwell's new law was that a changing electric field creates a magnetic field (the principle behind the electric motor). This was a symmetrical counterpart to Faraday's law of induction which states that a changing magnetic field creates an electric field (the principle behind the electric generator).

[4] Even though I promised an equation-free book, for those who are curious, I will slip them into this footnote where no one else will see them:

$\textit{div}\ \mathbf{E} = \rho$

$\textit{div}\ \mathbf{B} = 0$

$\textit{curl}\ \mathbf{E} = -\ \textit{chng}\ \mathbf{B}/c$

$\textit{curl}\ \mathbf{B} = (\textit{chng}\ \mathbf{E} + \mathbf{j})/c.$

E and **B** are the electric and magnetic fields; the use of bold font indicates that the fields have a direction. ρ is the density of electric charge, **j** is the electric current density, and *c* is a constant equal to the speed of light. The terms *div* and *curl* are mathematical terms that describe how a field at a point changes spatially in an outward direction (divergence) and in a rotary way around the point (curl). *Chng* is my "pidgin-math" for rate of change - what mathematicians call the time derivative.

[5] Because of Hertz's Jewish ancestry, the Nazis renamed the unit after Hermann von Helmholtz, whose name could also be abbreviated Hz. I've been told that one professor said to his class: "The new unit of frequency is to be the Helmholtz, abbreviated Hz and pronounced *Hertz.*"

[6] Wavelength is the distance between peaks of field intensity; it can be calculated by dividing *c* (the speed of light) by the frequency. The shortest wavelength of microwaves is about one millimeter, so the term "micro" is really a misnomer.

[7] Planck himself, although antagonistic to Hitler, remained in Germany throughout the war. He died in 1947 as plans were being made to honor him on his 90th birthday.

[8] The other two constants predicted by Planck's theory are Avogadro's number (that specifies the mass of atoms) and Boltzmann's constant (that relates our temperature scale to molecular energy).

[9] The use of the suffix "-on" for "particle" began with the word "electron", suggested by Irish physicist George Stoney in 1891.

[10] There is also a third kind of field that exists in space even if there are no quanta or self-fields. It is known as the vacuum field, described in Chap. 7.

Chapter 4

[1] Hero, who also built the first steam engine, found that water could not fill a vessel unless the air within it was allowed to escape, showing that air has substance. He also found that air remaining in a vessel could be squeezed into a smaller volume by increasing the water pressure. The only way this could happen, Hero reasoned, was if air consists of individual particles separated by space.

[2] The neutron also plays an important role in nuclear fission, which is the basis for nuclear power and for the atomic bomb. Because of its electrical neutrality, it is able to enter a nucleus and upset the balance (see "A delicate balance' below), causing the nucleus to break up and release energy.

[3] The reason that four nucleons escape as a unit is because they are bound

together so tightly. An analogy might be a group of four prisoners working together who have a better chance of escape than a single prisoner.

[4] Germany's atomic bomb project was under the direction of Werner Heisenberg, originator of the *uncertainty principle* of quantum mechanics. After the war, Heisenberg maintained that he deliberately slowed down the research because of moral scruples, but whether he did or not is, of course, uncertain.

[5] This can be seen mathematically, if you don't mind my slipping another equation into a footnote. Yukawa's equation, without including interactions, is $chng^2\ U/c^2 = Lap\ U - \lambda^2\ U$. Here U is the symbol he used for the field, $chng^2$ is my "pidgin math" for how the rate of change of the field changes with time (what mathematicians call the second time derivative), and *Lap* is a symbol for how the field values change from point to point (what mathematicians call the Laplacian). It is not hard to see how the time development (the $chng^2$ term) is impeded by the negative λ term.

Chapter 5

[1] The Conseil Européen pour la Recherche Nucléaire (CERN) is the world's largest particle physics laboratory, situated on the border between France and Switzerland. It is also known as the birthplace of the world wide web (www).

[2] This was a misunderstanding. The x-rays do not come from the "phosphorescent walls" of the Crookes tube; they are emitted from the metal plate inside the tube where the electron beam is brought to a screeching halt. But if not for this misunderstanding, Becquerel might not have tried his experiment with phosphorescent chemicals, and… who knows?

[3] Half-life is the time it takes for half of the radioactive atoms to decay. This does not mean that all of them have decayed after two half-lives. During the second half-life, half of the remaining nuclei decay, leaving one-quarter; after three half-lives there will be one-eighth left, etc.

[4] We have already encountered alpha rays (helium nuclei) in Chapter 4 and gamma rays (EM radiation) in Chapter 3.

[5] Wolfgang Pauli had similar feelings. When his wife left him for a chemist, he said in disbelief, "Had she taken a bullfighter I would have understood, but a *chemist*!" (B2003, p. 137)

[6] Another of Klein's accomplishments, along with Walter Gordon, was to derive the Yukawa equation before Yukawa did, but they derived it for the wrong particle (the electron). When the electron spin was discovered, the

equation was abandoned, only to be rediscovered by Yukawa, who applied it to the right particle — the pion. Nevertheless, the equation is known today as the Klein-Gordon equation.

[7] Another way to test for left-right symmetry is to imagine that you are communicating by radio with intelligent beings in another galaxy. If nature is left-right symmetric, there is no experiment that will enable them to determine what you mean by left and right.

[8] This nuclear magnetism is the basis for the medical procedure known as Magnetic Resonance Imaging.

[9] *T* stands for *tensor*, *V* for *vector* and *A* for *axial vector*, but you really don't need to know that, nor do you need to know what those terms mean. You don't even need to know that there might be a little *S* for *scalar* mixed in with the *T*. These are mathematical subtleties without any impact on the basic concepts.

[10] According to a later count, there were 13 wrong experiments (M2000, p. 414).

[11] A boson is a field quantum with integral spin, i.e., one of the four force fields.

Chapter 6

[1] The discovery that sunlight passed through a prism produces a color spectrum was made by Newton in 1665.

[2] Schrödinger, a Catholic, would not live in a country in which persecution of Jews became national policy. He eventually settled at the newly-created Institute of Advanced Studies in Dublin. "He did it, and we admired him. For it is no small matter to be uprooted in one's middle age, and to live in a foreign land. But he would have it no other way." – Max Born (B1966, p. 1066)

[3] The spin value of the neutron was determined by Julian Schwinger when he was a 19-year-old graduate student!

[4] The sound you hear in the background is Isaac Newton turning over in his grave (cf. quote at beginning of Chap. 2).

[5] Ironically, this account of how Quantum Field Theory was perfected was published in a collection entitled "The birth of *particle* physics".

Chapter 7

[1] Bell is best known for "Bell's Theorem", which was called the most

important contribution to quantum mechanics in the last 50 years. Unfortunately, his article "How to teach special relativity" is not well-known.

[2] In his later years Einstein maintained that he was unaware of Michelson's experiment when he wrote his 1905 paper (P1982, p. 115-117; I2007, p. 116-117). However, in the paper there is a clear reference to "the unsuccessful attempts to discover any motion of the earth relatively to the light medium".

[3] Except for dimensions, this is essentially the same apparatus that was used to detect gravity waves in 2016 (see Chapter 2).

[4] You can either take my word for this or work it out with some high school algebra.

[5] Einstein realized later that this "apparently irreconcilable" postulate was not necessary. The speed of light is dictated by Maxwell's equations, and if these equations are unchanged by motion then the velocity of light must be constant.

[6] There are more than 20,000 websites that deal with this paradox, and the Wikipedia article alone has over 3000 words.

[7] Of course disturbances that propagate in the backward direction have a shorter distance to travel, but if you work out the math, this effect is be not as great.

[8] Here is the calculation, tucked safely into a footnote. Let d be the distance between the two men (taking into account the Lorentz contraction), v the speed of the raft, and t the time for the voice to travel from A to B. The distance traveled by B during this time is vt and the total distance that the sound must travel is the hypotenuse of the triangle: $\sqrt{(d^2 + v^2t^2)}$. Now if the speed of sound is c, that distance must also equal ct. By equating the two expressions we find that $t = \sqrt{(d^2/(c^2 - v^2)}$. By comparing this with the time required if the platform were stationary, which is simply d/c, we find that the time when moving is longer by a factor $1/\sqrt{(1 - v^2/c^2)}$. Eureka! We have just derived a basic result of relativity theory using the bottom-up approach.

[9] "Suppose that a beacon is stationed on Earth, high atop Mount Olympus, and an identical one is mounted on the starship Argo before it lifts off. The beacons emit light pulses at a uniform interval, which is the period of these 'clocks'. The epic journey begins. The once and future Jason, on the Argo, watches the Earth beacon, while Zeus, atop the mountain, looks up at the flashing light in the heavens. On the outward journey, each sees the other rapidly moving away, at speed v. This motion has two effects. Time runs more slowly on the relatively moving body; each successively received pulse

is emitted from a greater distance and takes an additional amount of time to reach the viewer. So each sees the other's clock running more slowly. This continues until the Argo reaches Vega. Jason judges that trip to take considerably less time than does Zeus, because, to Jason, the distance traversed is considerably smaller. The Argo turns around. Now Jason is advancing to meet Zeus' light bolts; each successive one has a shorter distance to travel and arrives sooner, which overcompensates the slowing of time. So Jason sees Zeus' clock running faster, all the way home. To Zeus, almost the whole round-trip time has elapsed when he receives the last light pulse emitted by the Argo before it turned about and headed for home. That is because Argo, moving at almost the speed of light, is not far behind! All the pulses emitted on the return journey, when Jason's clock [beacon!] also is seen as running fast, are received in that short time interval… Although each viewer finds the other clock running more slowly on the outbound leg, Zeus has that experience for much longer than does Jason. And, on the return trip, although each sees the other's clock running faster, that episode is of very brief duration for Zeus. In short, Zeus sees the round trip taking much longer than does Jason." (S1986, p. 117)

[10] Einstein was wrong to say that Lorentz "assumed" the contraction effect. Lorentz had actually done an elaborate and difficult calculation to show that the mechanism of contraction is real.

[11] Lucretius had addressed the same question in 60 BC: "Time by itself does not exist; but from things themselves there results a sense of what has already taken place, what is now going on and what is to ensue. It must not be claimed that anyone can sense time by itself apart from the movement of things." (B1966, p. 12)

Chapter 8

[1] i is defined as the square root of -1.

Chapter 10

[1] As we will see later, the discovery of dark matter may add an eighth field to the mix.

[2] It is straightforward math to show that the frequency of oscillation is given by $f = mc^2/h$, where h is Planck's constant.

[3] Planck's Law says that the energy of a quantum is given by $e = hf$, where f is frequency and h is Planck's constant. Combining this with the equation in footnote 3 gives $e = mc^2$.

[4] The reader should be warned that many physicists do not accept quantum collapse. They believe that a superposition of states continues until a later time. For example, Roger Penrose suggested that collapse occurs when the gravitational energy exceeds a certain amount (P1994, p. 339).

[5] Zweig had used the word *ace* to describe these quanta, but Gell-Mann's "quark" prevailed.

[6] S. Aoki et al., "Quenched Light Hadron Spectrum", Phys Rev Lett. 2000 Jan 10; 84(2):p. 238-41.

[7] The planets Neptune and Pluto were discovered for the same reason.

[8] The advanced reader, knowing that "at the same time" depends on your state of motion, may ask which reference frame is being referred to. Although there is no experimental evidence, I would suggest that the frame of the absorbing atom is appropriate, or at least a good possibility.

Contents