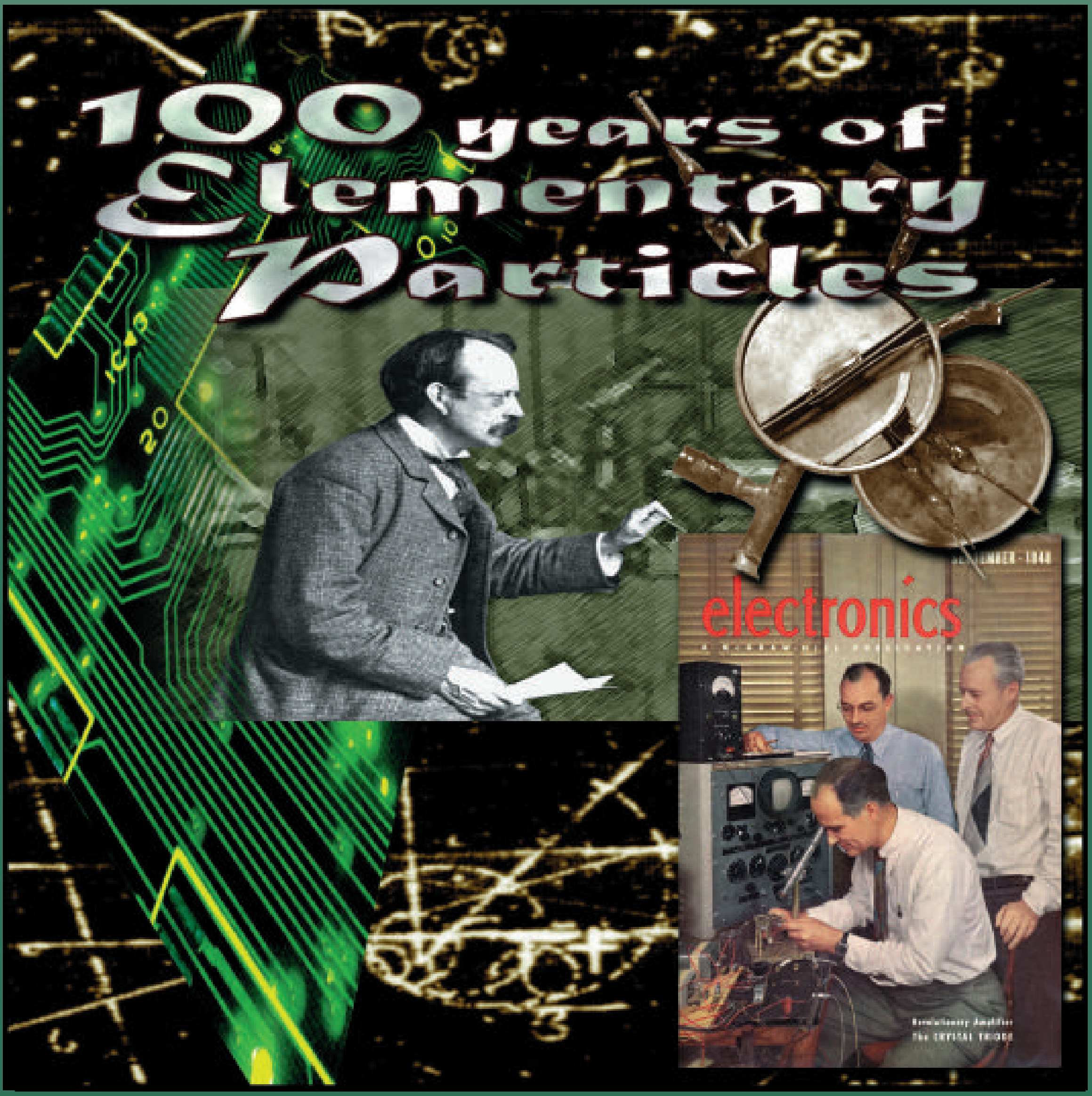
STANFORD LINEAR ACCELERATOR CENTER

Beam Line

Spring 1997, Vol. 27, No.1



Beam Line

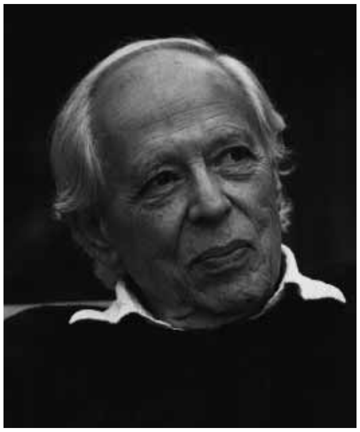
A PERIODICAL OF PARTICLE PHYSICS

SPRING 1997 VOL. 27, NUMBER 1

The *Beam Line* is published by the
Stanford Linear Accelerator Center,
PO Box 4349, Stanford, CA 94309.
Telephone: (415) 926-2585
INTERNET: beamline@slac.stanford.edu
FAX: (415) 926-4500
Issues of the *Beam Line* are accessible electronically on the World Wide Web at http://www.slac.stanford.edu/pubs/beamline
SLAC is operated by Stanford University under contract with the U.S. Department of Energy. The opinions of the authors do not necessarily reflect the policy of the Stanford Linear Accelerator Center.

Cover: A montage of images created by Terry Anderson for this special Beam Line *issue on the centennial of the electron's discovery. (Photograph of* Electronics *magazine courtesy AT&T Archives.)*

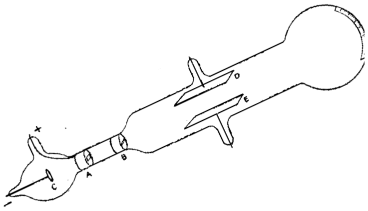
Printed on recycled paper



Pais

Weinberg

FEATURES



4 THE DISCOVERY OF THE ELECTRON
Abraham Pais

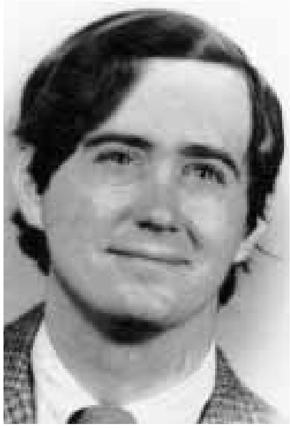
17 WHAT IS AN ELEMENTARY PARTICLE?
Steven Weinberg

22 ELEMENTARY PARTICLES: YESTERDAY, TODAY, AND TOMORROW
Chris Quigg

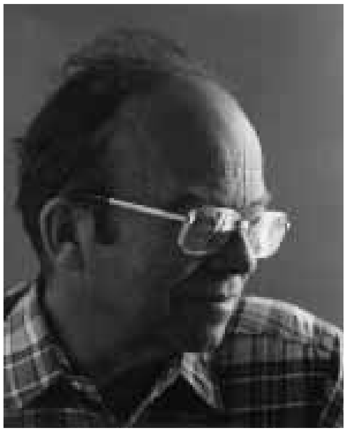
30 THE INDUSTRIAL STRENGTH PARTICLE
Michael Riordan

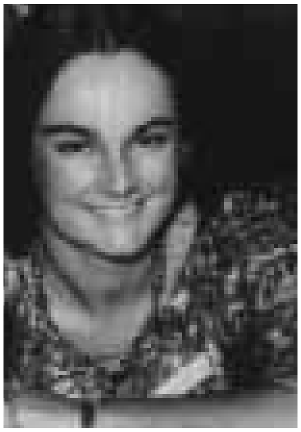
36 THE EVOLUTION OF PARTICLE ACCELERATORS & COLLIDERS
Wolfgang K. H. Panofsky

CONTENTS

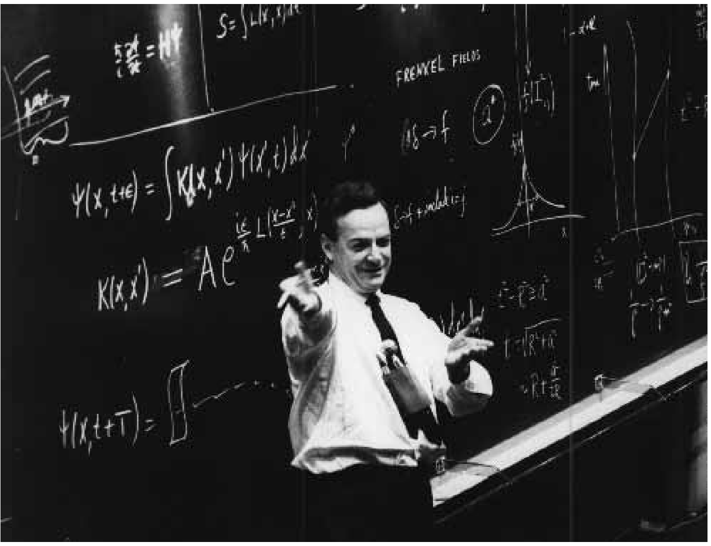

Quigg


Riordan


Panofsky


Trimble

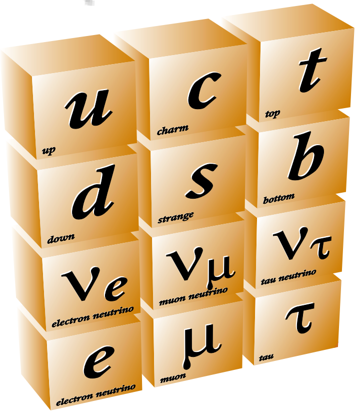
45 *THE UNIVERSE AT LARGE*:
THE ASTRO-PARTICLE-COSMO-CONNECTION
Virginia Trimble

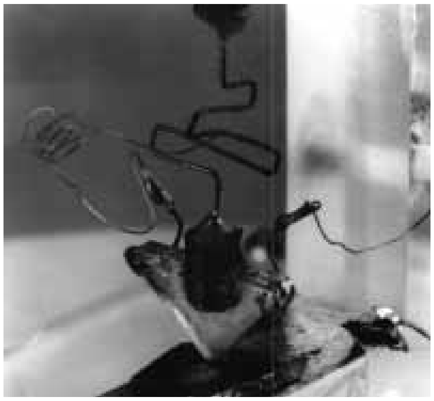


DEPARTMENTS

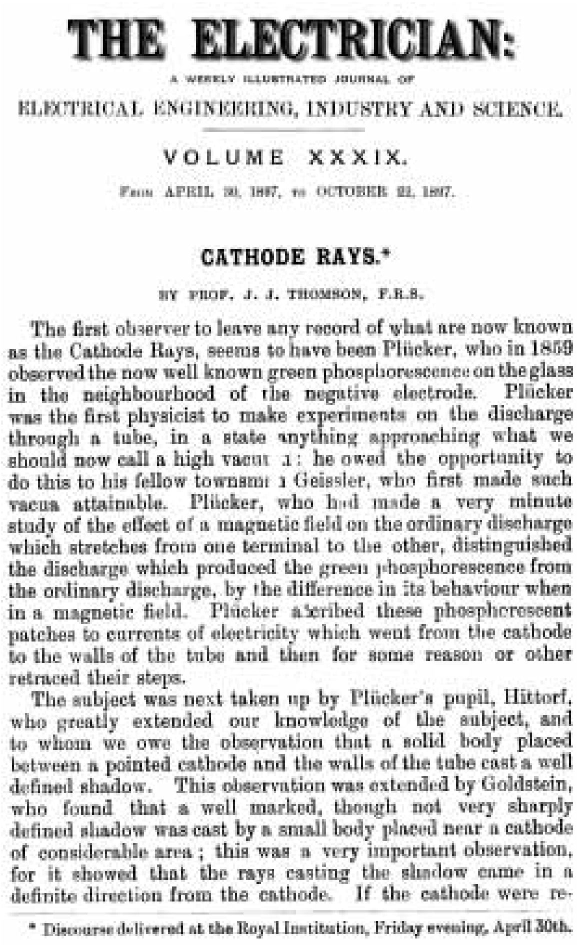
2 FROM THE EDITOR'S DESK

52 CONTRIBUTORS





FROM THE EDITOR'S DESK



THE ELECTRICIAN:

A WEEKLY ILLUSTRATED JOURNAL OF
ELECTRICAL ENGINEERING, INDUSTRY AND SCIENCE.

VOLUME XXXIX.

From APRIL 30, 1897, to OCTOBER 22, 1897.

CATHODE RAYS.*

BY PROF. J. J. THOMSON, F.R.S.

The first observer to leave any record of what are now known as the Cathode Rays, seems to have been Plücker, who in 1859 observed the now well known green phosphorescence on the glass in the neighbourhood of the negative electrode. Plücker was the first physicist to make experiments on the discharge through a tube, in a state anything approaching what we should now call a high vacuum: he owed the opportunity to do this to his fellow townsman Geissler, who first made such vacua attainable. Plücker, who had made a very minute study of the effect of a magnetic field on the ordinary discharge which stretches from one terminal to the other, distinguished the discharge which produced the green phosphorescence from the ordinary discharge, by the difference in its behaviour when in a magnetic field. Plücker ascribed these phosphorescent patches to currents of electricity which went from the cathode to the walls of the tube and then for some reason or other retraced their steps.

The subject was next taken up by Plücker's pupil, Hittorf, who greatly extended our knowledge of the subject, and to whom we owe the observation that a solid body placed between a pointed cathode and the walls of the tube cast a well defined shadow. This observation was extended by Goldstein, who found that a well marked, though not very sharply defined shadow was cast by a small body placed near a cathode of considerable area; this was a very important observation, for it showed that the rays casting the shadow came in a definite direction from the cathode. If the cathode were re-

---

* Discourse delivered at the Royal Institution, Friday evening, April 30th.

Introductory paragraphs from Thomson's paper, as published in the May 21, 1897 issue of The Electrician.

THE ELECTRON—or at least our recognition of its existence as an elementary particle—passes the century mark this spring. On April 30, 1897, Joseph John Thomson reported the results of his recent experiments on cathode rays to a Friday evening meeting of the Royal Institution, suggesting these rays were composed of negatively charged pieces of atoms that he dubbed "corpuscles." Six months later he published an extensive account of these experiments in the *Philosophical Magazine.* One of the classic papers of modern physics, it opened the doors of human consciousness to a radically new and often baffling world within atoms, one that has provided fertile ground for much of twentieth-century physics.

Together with the discovery of X rays and radioactivity during the preceding two years, and the introduction of the quantum three years later, this breakthrough led to a revolutionary conception of matter that has since had major impacts on other sciences, on modern technology and art, and even on the way we talk and think. The smooth, continuous, comfortable world of late nineteenth-century Europe was shattered into myriad bewildering fragments—some of which interact via forces that nobody had ever before encountered. Whether atoms themselves existed or not was in hot dispute at the time; among those who believed they did were prominent physicists who regarded them as vortex rings in the luminiferous aether. A century later, despite many superb advances, we are still struggling to achieve a grand synthesis of all the disparate shards encountered since.

To commemorate this pivotal breakthrough—and, in a more catholic sense, the discovery of

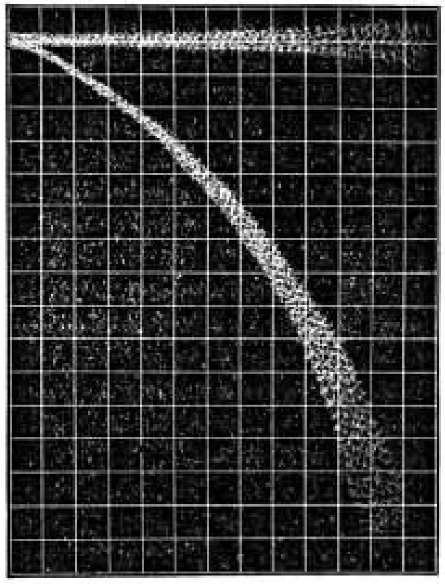


FIG. 5.—Hydrogen (Ammeter, 12 ; Voltmeter, 1,600).

Illustration from Thomson's article showing luminous paths of cathode rays (lower trace) bending in a magnetic field. The upper trace is due to ionized atoms in the gas.

subatomic particles—the *Beam Line* Editorial Board organized this special anniversary issue and asked me to serve as its guest editor. It has been a truly stimulating and rewarding experience. I am privileged to have worked with some of the nation's most literate physicists, who have contributed perceptive essays in honor of Thomson's fabulous discovery.

Three theorists open this issue by offering us their perspectives on the discovery, the meaning and the evolution of elementary particles. While Abraham Pais relates how the concept of the electron emerged from nineteenth-century research on electrochemistry and vacuum-tube discharges, Steven Weinberg and Chris Quigg take more modern and personal viewpoints. They examine what it means to call a particle "elementary" and try to assess where our discipline is headed as its second century begins.

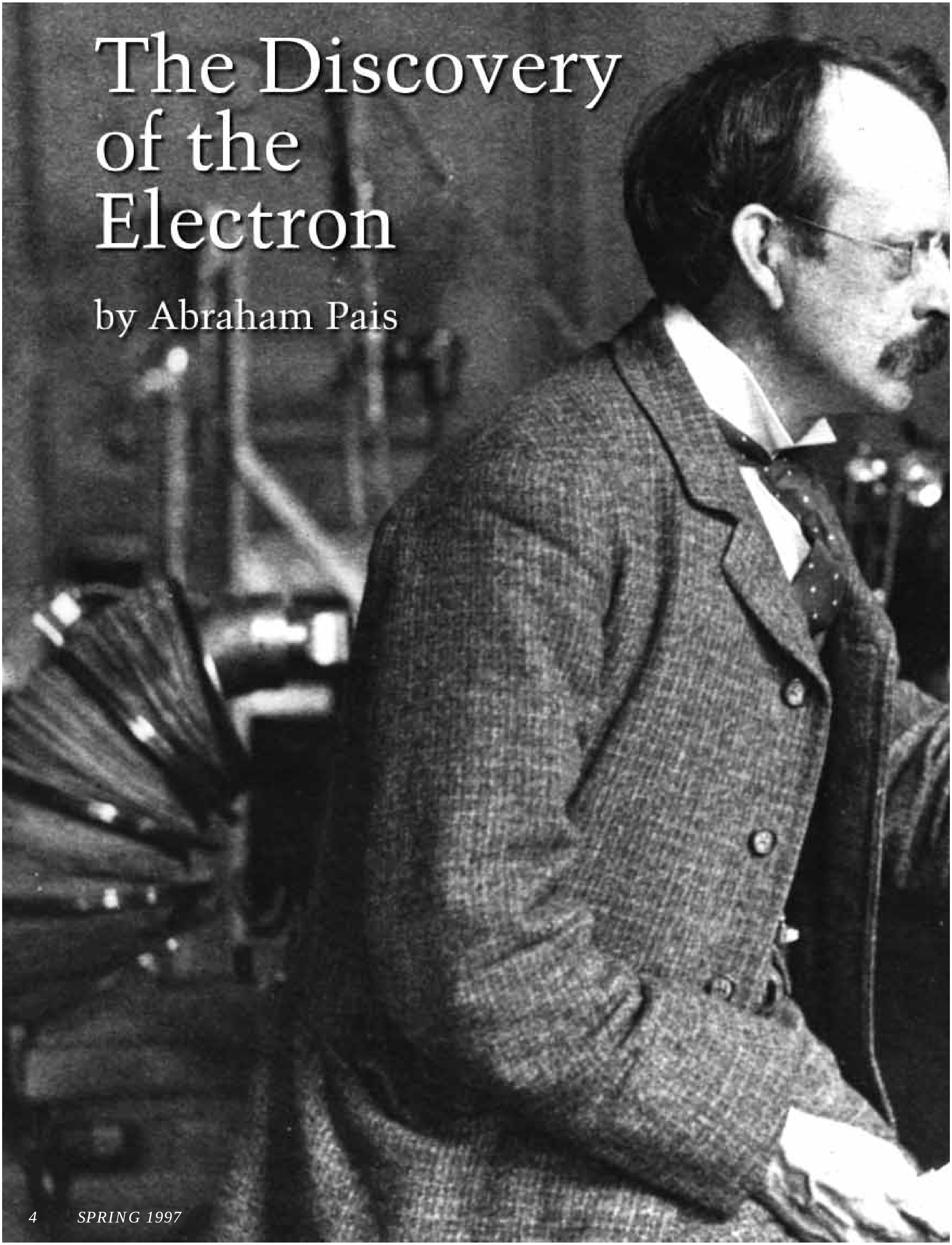
The final three articles concern "applications" of our knowledge of subatomic particles—in electronics technology, in pushing back the frontiers of high-energy research itself, and in understanding the origin and evolution of the Universe. My article indicates how our knowledge of the electron *as a particle* has proved crucial to the surging growth of what is now the world's biggest industry. Taking a retrospective look at particle accelerators and colliders, Wolfgang Panofsky evaluates various avenues being considered for the future of this technology. And Virginia Trimble closes this anniversary issue by surveying how the tiniest things in existence are closely linked to the structure and behavior of the largest.

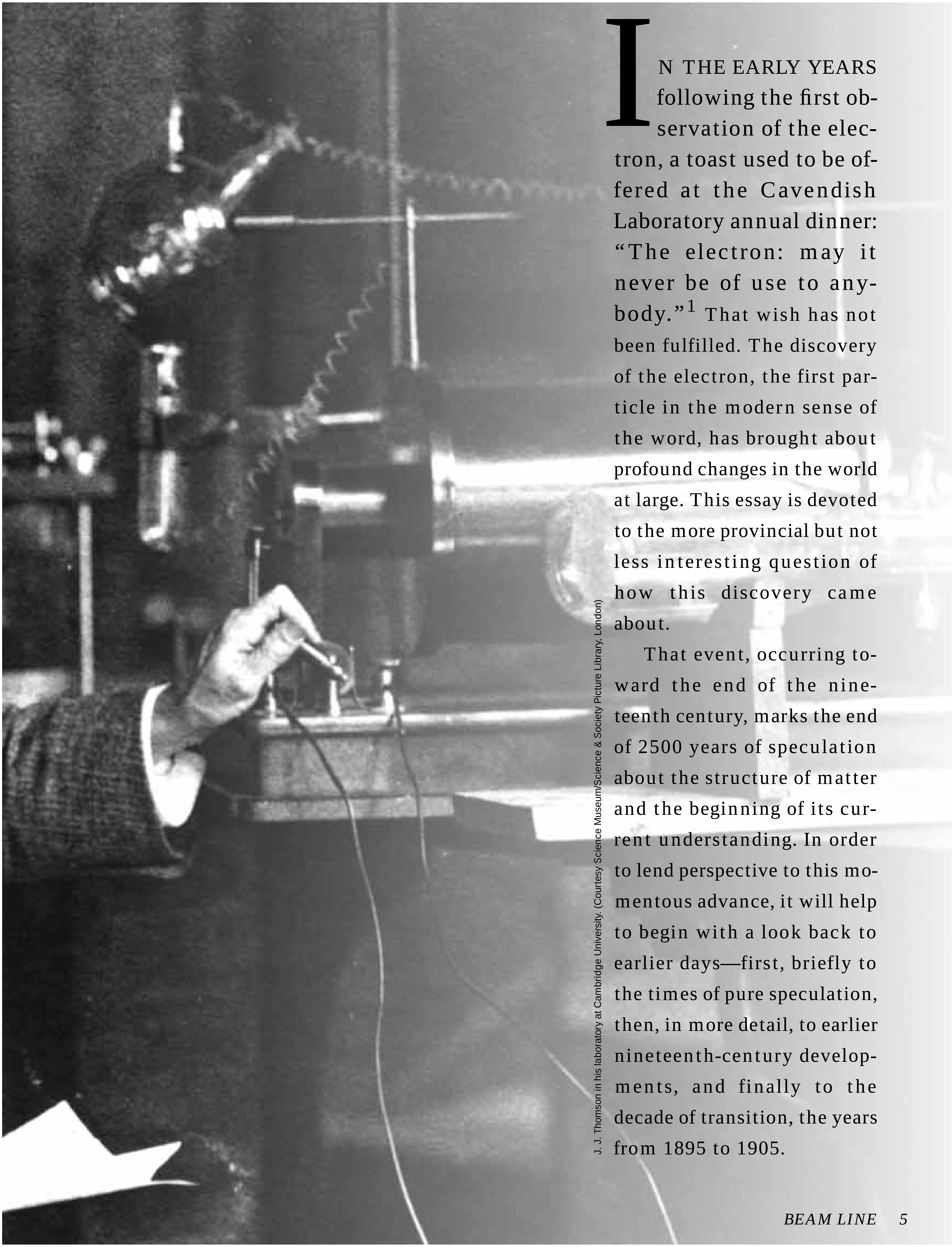
What will historians think, a hundred years hence, when they gaze back upon our own time? What conceptions that we hold dear today will be regarded then as we now regard the aether of 1897? What will be the "elementary particles" of the late twenty-first century? We can only guess. Whatever the answers, however, there can be little doubt that the hundred years that began with Thomson's discovery will be viewed as a remarkable period of scientific, technological and cultural achievement.

Michael Riordan

The Discovery of the Electron

by Abraham Pais





J. J. Thomson in his laboratory at Cambridge University. (Courtesy Science Museum/Science & Society Picture Library, London)

IN THE EARLY YEARS following the first observation of the electron, a toast used to be offered at the Cavendish Laboratory annual dinner: "The electron: may it never be of use to anybody."[1] That wish has not been fulfilled. The discovery of the electron, the first particle in the modern sense of the word, has brought about profound changes in the world at large. This essay is devoted to the more provincial but not less interesting question of how this discovery came about.

That event, occurring toward the end of the nineteenth century, marks the end of 2500 years of speculation about the structure of matter and the beginning of its current understanding. In order to lend perspective to this momentous advance, it will help to begin with a look back to earlier days—first, briefly to the times of pure speculation, then, in more detail, to earlier nineteenth-century developments, and finally to the decade of transition, the years from 1895 to 1905.

THE ANCIENTS

THE TERM *atom*, derived from the Greek α, a privative, and *τεμειν*, to cut, appears first, I am told, in the writings of Greek philosophers of the fifth century BC. Democritus (late fifth century BC) taught that atoms are the smallest parts of matter, though in his view they were not necessarily minute. Empedocles (490–430 BC), physicist, physician, and statesman, held that there are four indestructible and unchangeable elements—fire, air, water and earth—eternally brought into union and eternally parted from each other by two divine forces, love and discord. Nothing new comes or can come into being. The only changes that can occur are those in the juxtaposition of element with element. Epicurus' (341–270 BC) opinion that atoms cannot be divided into smaller parts by physical means, yet that they have structure, was shared by prominent scientists well into the nineteenth century AD. The Roman poet Lucretius (98–55 BC) was an eloquent exponent of the theory that atoms, infinite in number but limited in their varieties, are, along with empty space, the only eternal and immutable entities of which our physical world is made. Today's scientist will not fail to note that in each of these speculative thinkers' considerations one finds elements that sound curiously modern.

The opposite position, that matter is infinitely divisible and continuous, likewise had its early distinguished proponents, notably Anaxagoras (c 500–428 BC) and Aristotle (384–322 BC). The latter's prestige eclipsed the atomists' view until the seventeenth century. Even that late, Rene Descartes (1596–1650) pronounced that there cannot exist any atoms or parts of matter that are of their own nature indivisible; for though God had rendered a particle so small that it was not in the power of any creature to divide it, He could not, however, deprive Himself of the ability to do so.[2]

THE NINETEENTH CENTURY

REGARDING THE UNDERSTANDING of the basic structure of matter, very little had changed between the days of speculation by the ancient Greek philosophers and the beginning of the nineteenth century, when, in 1808, the British chemist and physicist John Dalton (1766–1844) commenced publication of his *New System of Chemical*

Philosophy. He had of course illustrious precursors, notably Antoine-Laurent Lavoisier (1743–1794). Yet his quantitative theory suddenly could explain or predict such a wealth of facts that he may properly be regarded as the founder of modern chemistry. In a sequel volume Dalton expressed the fundamental principle of the youngest of the sciences in these words:

> I should apprehend there are a considerable number of what may be properly called elementary principles, which can never be metamorphosed, one into another, by any power we can control. We ought, however, to avail ourselves of every means to reduce the number of bodies or principles of this appearance as much as possible; and after all we may not know what elements are absolutely indecomposable, and what are refractory, because we do not know the proper means of their reduction. All *atoms of the same kind*, whether simple or compound, must necessarily be conceived to be alike in shape, weight, and every other particular.

John Dalton, whose New System of Chemical Philosophy *resurrected the atomic theory of matter. (Courtesy A. L. Smyth,* John Dalton: 1766–1844, a Bibliography of Works By and About Him *and AIP Emilio Segrè Visual Archives)*

These superb lines ushered in the intense nineteenth century discussions on the nature of atoms and molecules. Perhaps the most remarkable fact about these debates is the great extent to which chemists and physicists spoke at cross purposes when they did not actually ignore each other. This is not to say that there existed one common view among chemists, another among physicists. Rather, in either camp there were many and often strongly diverging opinions. The principal point of debate among chemists was whether atoms were real objects or only mnemonic devices for coding chemical regularities and laws. The main issues for the physicists centered around the kinetic theory of gases, in particular around the meaning of the second law of thermodynamics.

An early illustration of the dichotomies between chemists and physicists is provided by the fact that Dalton did not accept the hypothesis put forward in 1811 by Amadeo Avogadro (1776–1856) that, for fixed temperature and pressure, equal volumes of gases contain equal numbers of molecules. Nor was Dalton's position held only by a single person for a brief time. The tardiness with which Avogadro's law came to be accepted clearly indicates the widespread resistance to the idea of molecular reality. As but one further illustration of this attitude I mention some revealing remarks by

Alexander Williamson (1824–1904), himself a convinced atomist. In his presidential address of 1869 to the London Chemical Society, he said:

> It sometimes happens that chemists of high authority refer publicly to the atomic theory as something they would be glad to dispense with, and which they are ashamed of using. They seem to look upon it as something distinct from the general facts of chemistry, and something which the science would gain by throwing off entirely. . . . On the one hand, all chemists use the atomic theory, and . . . on the other hand, a considerable number view it with mistrust, some with positive dislike.[3]

On the whole, molecular reality met with less early resistance in physics than it did in chemistry. That is not surprising. Physicists could already *do* things with molecules and atoms at a time when chemists could, for most purposes, take them to be real or leave them as coding devices.

The insight that gases are composed of discrete particles dates back at least to the eighteenth century. Daniel Bernoulli (1700–1782) may have been the first to state that gas pressure is caused by the collisions of particles with the walls that contain them. The nineteenth-century masters of kinetic theory were atomists—by definition, one might say. In Rudolf Clausius' (1822–1888) paper of 1857, entitled "On the kind of motion we call heat," the distinction between solids, liquids, and gases is related to different types of molecular motion. Ludwig Boltzmann (1844–1906) was less emphatic, but he could hardly have developed his theory of the second law of thermodynamics had he not believed in the particulate structure of matter.

Long before these learned fin-de-siecle discourses took place, in fact long before the laws of thermodynamics were formulated, theoretical attempts had begun to estimate the dimensions of molecules. As early as 1816, Thomas Young (1773–1829) noted that "the diameter or distance of the particles of water is between the two thousand and the ten thousand millionth of an inch."[4] In 1873 James Clerk Maxwell stated that the diameter of a hydrogen molecule is about 6×10^{-8} cm.[5] That same year Johannes Diderik van der Waals (1837–1923) reported similar results in his doctoral thesis.[6] By 1890 the spread in these values, and those obtained by others, had narrowed considerably. A review of the results up to the late 1880s placed the radii of hydrogen and air molecules between 1 and 2×10^{-8} cm,[7] a remarkably sensible range.

Until the very last years of the nineteenth century, most if not all scientists who believed in the reality of atoms shared the view that these particles cannot be decomposed further, as was eloquently expressed by Maxwell in 1873:

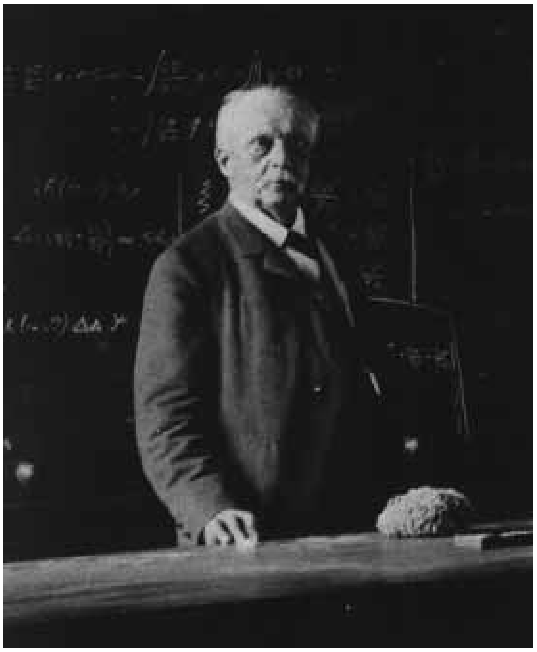
> Though in the course of ages catastrophes have occurred and may yet occur in the heavens, though ancient systems may be dissolved and new systems evolved out of their ruins, the molecules [i.e., atoms!] out of which these systems [the Earth and the whole solar system] are built—the foundation stones of the material universe—remain unbroken and unworn. They continue this day as they were created—perfect in number and measure and weight.[8]

THE ATOMICITY OF CHARGE

ELECTROMAGNETISM BECAME A PART of science in the eighteenth century, largely due to rapid progress in the invention of new instruments: the first condenser (the Leiden jar), the lightning rod, the first battery (the Voltaic pile), the first solenoid. These advances led, in turn, to the formulation of phenomenological laws based on new experiments. Of interest here is the law of electrolysis, formulated in the 1830s by Michael Faraday (1791–1867), one of the great experimentalists of all time, who coined terms of lasting use: electrode, anode, cathode, electrolysis, ion, anion, cation. In modern language, his law can be stated like this:

The amount of electricity deposited at the anode by a gram mole of monovalent ions is a universal constant, the farad (F), given by $F = Ne$, where N, Avogadro's number, is the number of molecules per mole, and e is a universal unit of charge.

What does this e signify? In 1881 Herman von Helmholtz (1821–1894) put it like this in his Faraday lecture: "The most startling result of Faraday's law is perhaps this. If we accept the hypothesis that the elementary substances are composed of atoms, we cannot avoid concluding that electricity also, positive as well as negative, is divided into definite elementary portions, which behave like atoms of electricity."[9] This statement explains why in subsequent years the quantity e was occasionally referred to in the German literature as "das Helmholtzsche Elementarquantum."



Hermann von Helmholtz, who in 1881 speculated on the atomicity of charge. (Courtesy AIP Emilio Segrè Visual Archives)

March of Discovery

1895

- Wilhelm Roentgen (1845–1923) discovers X rays, for which he would receive the first Nobel Prize in physics, in 1901.

1896

- Antoine Henri Becquerel (1852–1908) observes what he called “uranic rays,” the first phenomenon that opens a new field later called radioactivity.
- Wilhelm Wien (1864–1928) publishes his exponential law for black-body radiation, the first quantum law ever written down.
- Pieter Zeeman’s (1865–1934) first paper appears on the influence of magnetic fields on spectral lines.

1897

- Determination of e/m for cathode rays by J. J. Thomson and others.
- First mention of a particle lighter than hydrogen.

1898

- Ernest Rutherford discovers there are two species of radioactive radiations: α-rays and β-rays.

1899

- Thomson measures the electric charge of free electrons and realizes that atoms are split in ionization processes.

1900

- Paul Villard (1860–1934) discovers γ-rays.
- First determination of a half-life for radioactive decay.
- Max Planck discovers the quantum theory.

1905

- Albert Einstein postulates the light quantum (March).
- Einstein’s first paper on special relativity is published (June).

Even before Helmholtz’s memorable address, the Irish physicist George Johnstone Stoney (1826–1911) had reported to the 1874 meeting of the British Association for the Advancement of Science an estimate of e, the first of its kind, based on $F = Ne$. Values for F and N were reasonably well known by then. Stoney obtained $e \sim 3 \times 10^{-11}$ esu, too small by a factor of about 20, yet not all that bad for a first and very early try.[10] In 1891 he baptized the fundamental unit of charge, giving it the name “electron.”[11] Thus the term was coined prior to the discovery of the quantum of electricity and matter that now goes by this name.

DECADE OF TRANSITION

IN MARCH 1905 Ernest Rutherford (1871–1937) delivered the Silliman lectures at Yale. He began the first of his talks as follows:

> The last decade has been a very fruitful period in physical science, and discoveries of the most striking interest and importance have followed one another in rapid succession. . . . The march of discovery has been so rapid that it has been difficult even for those directly engaged in the investigations to grasp at once the full significance of the facts that have been brought to light. . . . The rapidity of this advance has seldom, if ever, been equaled in the history of science.[12]

The speed with which one important discovery followed another (see box at left) was indeed breathtaking. It is natural to ask but not easy to answer why so much novelty should be discovered in so short a time span. It is clear, however, that a culmination of advances in instrumentation was crucial. They include:

- *Higher voltages*. Higher voltages were the result of Heinrich Ruhmkoff’s (1803–1874) work, beginning in the 1850s, on an improved version of the induction coil. These were the coils that in 1860 served Gustav Kirchhoff (1824–1887) and Robert Bunsen (1811–1899) in their analysis of spark spectra; Heinrich Hertz (1857–1894) in 1886–1888 in his demonstration of electromagnetic waves and his discovery of the photoelectric effect; Wilhelm Roentgen in his discovery of X rays; Guglielmo Marconi (1874–1937) in his transmission of telegraph signals without wires; Pieter Zeeman in his discovery of the Zeeman effect; and Thomson in his determination of e/m for electrons. By the turn of the century, voltages of the order of 100,000 volts could be generated by these coils.

J. J. Thomson and Ernest Rutherford (right) at the Cavendish Lab in 1934. (Courtesy AIP Emilio Segrè Visual Archives Bainbridge Collection)

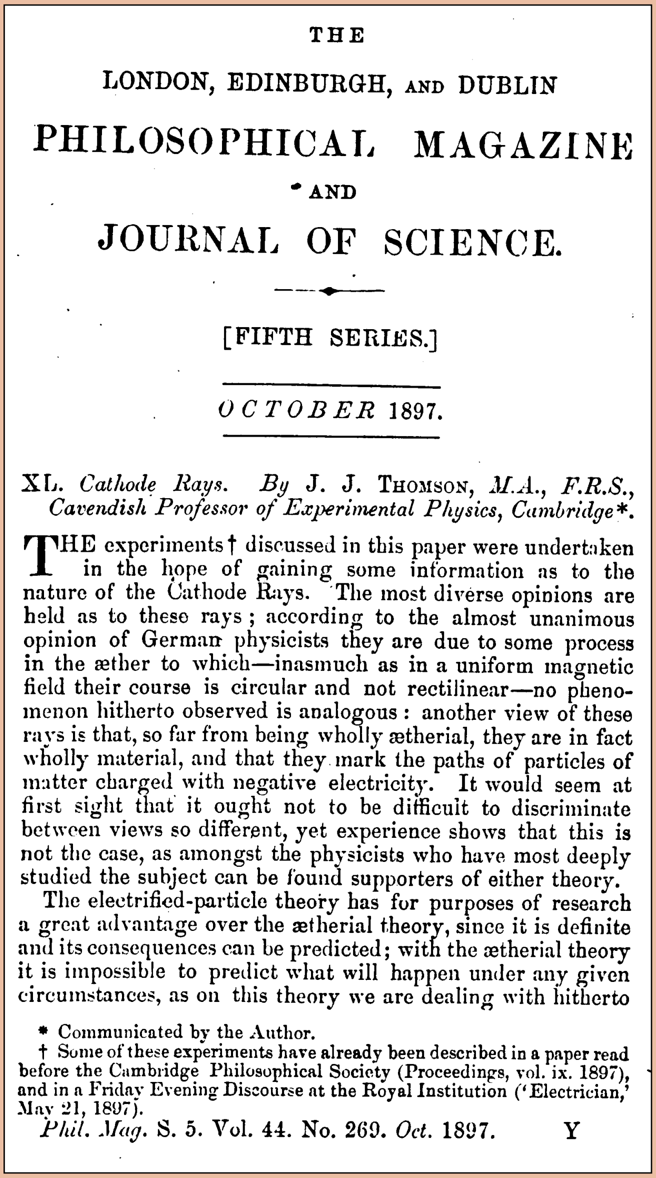
- *Improved vacua.* Improved vacua were achieved in the 1850s, when Johann Geissler (1815–1879) began developing the tubes now named after him. Soon he was able to reach and maintain pressures of 0.1 mm of mercury. Refined versions of this tube were crucial to the discoveries of Roentgen and Thomson.
- *Ionization chambers.* Early versions of the parallel-plate ionization chamber were developed in Cambridge during the 1890s. They were used by Rutherford and the Curies in the earliest quantitative measurements of radioactivity.
- *Concave spectral gratings.* Concave spectral gratings were developed starting in the 1880s by Henry Rowland (1848–1901) at the Johns Hopkins University. Their resolving power made Zeeman's discovery possible.
- *Cloud chambers.* Work on the development of a cloud chamber was begun in Cambridge in 1895 by Charles T. R. Wilson (1869–1959). This instrument enabled Thomson to measure the electron's charge.

THE DISCOVERY

ALL RESEARCH THAT LED to the discovery of the electron deals with studies of cathode rays, a subject that had already engaged Faraday, who in 1838 made this prophetic remark on its future: "The results connected with the different conditions of positive and negative discharge will have a far greater influence on the philosophy of electrical science than we at present imagine."[13]

J. J. Thomson discovered the electron. Numerous are the books and articles in which one finds it said that he did so in 1897. I cannot quite agree. It is true that in that year Thomson made a good determination of e/m for cathode rays, an indispensable step toward the identification of the electron, but he was not the only one to do so. It is also true that in 1897 Thomson correctly conjectured that the large value for e/m he had measured indicated the existence

Thomson's Two Experimental Papers



THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[FIFTH SERIES.]

OCTOBER 1897.

XL. *Cathode Rays*. *By* J. J. THOMSON, *M.A.*, *F.R.S.*, *Cavendish Professor of Experimental Physics, Cambridge*[*].

THE experiments[†] discussed in this paper were undertaken in the hope of gaining some information as to the nature of the Cathode Rays. The most diverse opinions are held as to these rays; according to the almost unanimous opinion of German physicists they are due to some process in the æther to which—inasmuch as in a uniform magnetic field their course is circular and not rectilinear—no phenomenon hitherto observed is analogous: another view of these rays is that, so far from being wholly ætherial, they are in fact wholly material, and that they mark the paths of particles of matter charged with negative electricity. It would seem at first sight that it ought not to be difficult to discriminate between views so different, yet experience shows that this is not the case, as amongst the physicists who have most deeply studied the subject can be found supporters of either theory.

The electrified-particle theory has for purposes of research a great advantage over the ætherial theory, since it is definite and its consequences can be predicted; with the ætherial theory it is impossible to predict what will happen under any given circumstances, as on this theory we are dealing with hitherto

[*] Communicated by the Author.

[†] Some of these experiments have already been described in a paper read before the Cambridge Philosophical Society (Proceedings, vol. ix. 1897), and in a Friday Evening Discourse at the Royal Institution ('Electrician,' May 21, 1897).

Phil. Mag. S. 5. Vol. 44. No. 269. *Oct.* 1897. Y

WHEN J. J. THOMSON began his research on the cathode rays during the 1890s, there was great confusion about their exact nature. As he noted in the introduction to his paper, "On Cathode Rays," [*Phil. Mag.*, Ser. 5, Vol. 44, No. 269 (1897), p. 293]:

> The most diverse opinions are held as to these rays; according to the almost unanimous opinion of German physicists they are due to some process in the æther to which . . . no phenomenon hitherto observed is analogous; another view of these rays is that, so far from being wholly ætherial, they are in fact wholly material, and that they mark the paths of particles of matter charged with negative electricity.

Following the lead of French physicist Jean Perrin, Thomson first satisfied himself that the rays were negatively charged, then addressed a quandary that had been puzzling scientists on both sides of the Channel for years. Although the rays were easily deflected by a magnetic field, they were apparently *not* deflected by an electric field between two plates. The absence of this deflection, he showed, was due to the ionization of the gas remaining in a cathode-ray tube, which permitted a current to flow between the plates and drastically reduced the field. This did not occur at high vacuum, however, and the rays were indeed deflected as expected for negatively charged particles. Thus he noted:

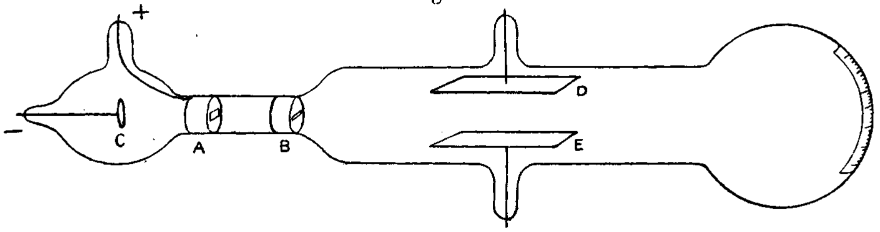
> I can see no escape from the conclusion that they are charges of negative electricity carried by particles of matter. The question next arises, What are these particles? [A]re they atoms, or molecules, or matter in a still finer state of subdivision?

By simultaneously deflecting the rays in *both* electric and magnetic fields, Thomson was able to determine their velocity and the ratio m/e of the mass m to the electric charge e carried by these (then) hypothetical particles. His result was startling:

> From these determinations we see that the value of m/e is independent of the nature of the gas, and that its value 10^{-7} [gram per emu] is very small compared with the value 10^{-4}, which is the smallest value of this quantity previously known, and which is the value for the hydrogen ion in electrolysis.

But he could not conclude from these data that m itself therefore had to be very small. "The smallness of m/e may be due to the smallness of m or the largeness of e," Thomson wrote. Because the values of m/e were independent of the nature and pressure of the gas, he began

Fig. 2.



The rays from the cathode C pass through a slit in the anode A, which is a metal plug fitting tightly into the tube and connected with the earth; after passing through a second slit in another earth-connected metal plug B, they travel between two parallel aluminium plates about 5 cm. long by 2 broad and at a distance of 1·5 cm. apart; they then fall on the end of the tube and produce a narrow well-defined phosphorescent patch. A scale pasted on the outside of the tube serves to measure the deflexion of this patch.

Figure from Thomson's first paper (together with explanatory text) illustrating the apparatus he used to measure e/m.

to envision atoms as made of "primordial atoms, which we shall for brevity call corpuscles." He went on:

> The smallness of the value of *m/e* is, I think, due to the largeness of *e* as well as the smallness of *m*. There seems to me to be some evidence that the charges carried by the corpuscles in the atom are large compared with those carried by the ions of an electrolyte.

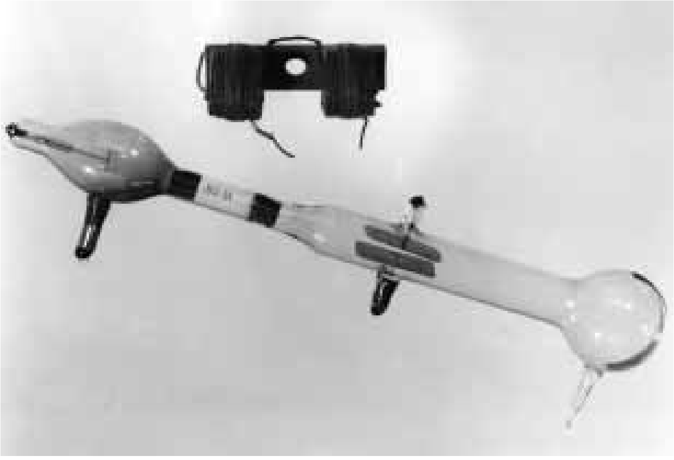
Over the next two years, Thomson determined the mass and charge of his corpuscles, but it took additional experiments culminating in a second paper, "On the Masses of the Ions in Gases at Low Pressures," [*Phil. Mag.,* Ser. 5, Vol. 48, No. 295 (1899), p. 547]. Using a novel technique developed by his student C. T. R. Wilson, he measured both *m/e* and *e* for the negatively charged particles created by dissociation of atoms in ultraviolet light. He found *m/e* to be the same as for cathode rays and *e* to have the same absolute value as the hydrogen ion in electrolysis. Thus he concluded:

> The experiments just described, taken in conjunction with the previous ones on the value of *m/e* for the cathode rays . . . show that in gases at low pressures negative electrification, though it may be produced by very different means, is made up of units each having a charge of electricity of a definite size; the magnitude of this negative charge is about 6 x 10^{-10} electrostatic units, and is equal to the positive charge carried by the hydrogen atom in the electrolysis of solutions.
>
> In gases at low pressures these units of negative electric charge are always associated with carriers of a definite mass. This mass is exceedingly small, being only about 1.4 x 10^{-3} of that of the hydrogen ion, the smallest mass hitherto recognized as capable of a separate existence. The production of negative electrification thus involves the splitting up of an atom, as from a collection of atoms something is detached whose mass is less than that of a single atom.

Thus was the first elementary particle finally discovered and the field of particle physics born. Educated at Cambridge as a mathematical physicist, Thomson seems to have grasped the importance of his breakthrough almost immediately. For he ended his second paper with some bold speculations about its ultimate significance:

> From what we have seen, this negative ion must be a quantity of fundamental importance in any theory of electrical action; indeed, it seems not improbable that it is the fundamental quantity in terms of which all electrical processes can be expressed. For, as we have seen, its mass and its charge are invariable, independent both of the processes by which the electrification is produced and of the gas from which the ions are set free. It thus possesses the characteristics of being a fundamental conception in electricity; and it seems desirable to adopt some view of electrical action which brings this conception into prominence.

Within a few years most physicists recognized Thomson's new particle by the name "electron," the term George Stoney had coined for the fundamental unit of charge (see main text). But Thomson stuck resolutely by his beloved "corpuscle" and still refused to call it anything else upon receiving the 1906 Nobel Prize in Physics "in recognition of the great merits of his theoretical and experimental investigations on the conduction of electricity by gases." —M.R.



The vacuum tube used by Thomson in his discovery of the electron. (Courtesy Science Museum/Science & Society Picture Library, London)

of a new particle with a very small mass on the atomic scale. However, he was not the first to make that guess. In order to explain, I need to introduce two other players in the field.

The first is Emil Wiechert (1861–1928), then a Privatdozent at the University of Konigsberg. In the course of a lecture before Konigsberg's Physical Economical Society, on January 7, 1897, he stated his conclusion about cathode rays[14] to which his experiments had led him: "It showed that we are not dealing with the atoms known from chemistry, because the mass of the moving particles turned out to be 2000–4000 times smaller than the one of hydrogen atoms, the lightest of the known chemical atoms." It was the first time ever that a subatomic particle is mentioned in print and sensible bounds for its mass are given. However, these conclusions depended crucially on his assumption about the charge. "Als Ladung ist 1 Elektron angenommen" (the charge is assumed to be one electron) he stated, using Stoney's terminology.

The second person is Walter Kaufmann (1871–1947), then Assistent at the University of Berlin, whose cathode-ray experiments had taught him two crucial points.[15] First, e/m for his rays was a *constant*, the same for whatever residual gas remained in his Geissler tube. That greatly puzzled him: "This assumption [of constant e/m] is physically hard to interpret; for if one makes the most plausible assumption that the moving particles are ions [in the electrolytic sense] then e/m should have a different value for each gas." Furthermore there was, as he perceived it, a second difficulty. Assuming e/m to be a constant, his measurements gave him about 10^7 emu/g for the value of e/m, "while for a hydrogen ion [e/m] equals only 10^4." Thus, he stated, "I believe to be justified in concluding that the hypothesis of cathode rays as emitted particles is by itself inadequate for a satisfactory explanation of the regularities I have observed."

Clearly Kaufmann was a fine experimentalist who, however, lacked the chutzpah of Thomson, who on August 7, 1897, submitted his memoir on cathode rays.[16] His first determination of e/m yielded a value 770 times that of hydrogen. He observed (see box on pages 12 and 13) that, “The smallness of m/e may be due to the smallness of m or the largeness of e, or to a combination of these two.” He went on to argue in favor of the smallness of m, “Thus on this view we have in the cathode rays matter in a new state, a state in which the subdivision of matter is carried very much further than in the ordinary gaseous state: a state in which all matter . . . is of one and the same kind; this matter being the substance from which all the chemical elements are built up.”

As I see it, Thomson’s finest hour as an experimentalist came in 1899 when he applied the methods just described to photoelectrically produced particles and concluded—he was the first to do so!—that these particles were electrons: “The value of m/e in the case of ultraviolet light . . . is the same as for cathode rays.”[17] In the same paper he announced his experimental results for the value of e, obtained by a method recently discovered by his student C. T. R. Wilson, who had found that charged particles can form nuclei around which supersaturated water vapor condenses. Thomson’s measurement of e is one of the earliest applications of this cloud-chamber technique. He determined the number of charged particles by droplet counting, and their overall charge by electrometric methods, arriving at $e \sim 6.8\times10^{-10}$ esu, a very respectable result in view of the novelty of the method. And that is why Thomson is the discoverer of the electron.

When Thomson addressed a joint meeting of British and French scientists in Dover in 1899, most doubts had been resolved. He quoted a mass of 3×10^{-26} g for the electron, the right order of magnitude. The atom had been split. “Electrification essentially involves the splitting up of the atom, a part of the mass of the atom getting free and becoming detached from the original atom.”[18]

ENVOI

TO DEFINE the “birth of an era” is perhaps best left for parlor games. Let me write of the birth of particle physics nevertheless, define it to take place in 1897, and appoint Wiechert, Kaufmann and Thomson as keepers at the gate. Their respective experimental arrangements are of comparable quality, their experimental results equally good. Kaufmann’s observation that certain properties of cathode rays are independent of the nature of the gas they traverse is, we would say, a clear indication the universality of the constitution of these rays. The value for e/m he obtained is a good one. Had he added one conjectural line to his paper, something like, "If we assume e to be the fundamental unit of charge identified in electrolysis, then cathode rays must be considered to be a new form of matter," he would have shared equal honors with Thomson for advances made in 1897. Perhaps the thought never struck him, perhaps it did but was rejected as too wild. Perhaps also the Berlin environment was not conducive to uttering speculations of this kind, as is evidenced by a recollection about the year 1897: "I heard John Zeleny say that he was in Berlin at that time, working in the laboratory of Warburg. When the discovery of the electron was announced, nobody in Berlin would believe in it."[19] It may not have been known at that time what went through Kaufmann's mind; it certainly is not known now.

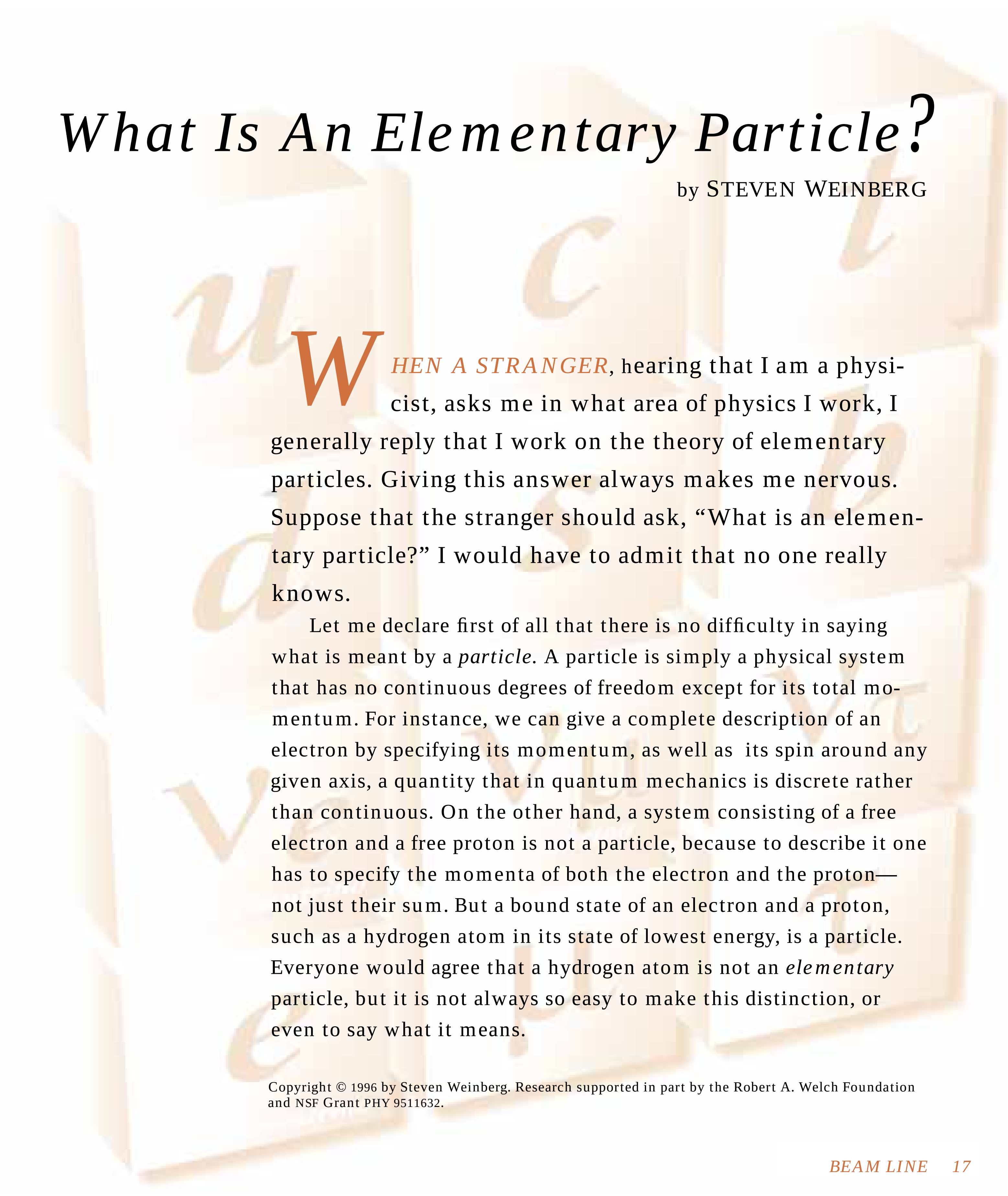
It is fitting to conclude with a line from one of my favorite essays: "On History," by Thomas Carlyle[20]: "No hammer in the Horologe of Time peals through the universe when there is a change from Era to Era. Men understand not what is among their hands."

NOTES

[1] E. N. da C. Andrade, *Rutherford and the Nature of the Atom* (New York: Doubleday, 1964), p. 48.
[2] R. Descartes, *Principles of Philosophy*, Part 2, Principle 20. Translated by E. Haldane and G. Ross (New York: Dover, 1955).
[3] A. Williamson, *J. Chem. Soc.* 22 (1869), 328.
[4] T. Young, *Miscellaneous Works*, Vol. 9, (New York: Johnson Reprint, 1972), p. 461.
[5] J. C. Maxwell, *Collected Works*, Vol. 2, (New York: Dover, 1952), p. 361.
[6] J. D. van der Waals, (Ph.D. diss., Sythoff, Leiden 1873).
[7] A. W. Rucker, *J. Chem. Soc.* 53 (1888), 222.
[8] Maxwell, *Collected Works*, pp. 376–77.
[9] H. von Helmholtz, in *Selected Writings by Hermann von Helmholtz*, ed. R.Kahl (Wesleyan Univ. Press, 1971), p. 409.
[10] G. J. Stoney, *Phil. Mag.* 11 (1881), 381.
[11] ———, *Trans. Roy. Dublin Soc.* 4 (1888–92), 563.
[12] E. Rutherford, *Radioactive Transformations* (London: Constable, 1906), pp. 1, 16.
[13] M. Faraday, *Philos. Trans. Roy. Soc.* 128 (1838), 125.
[14] E. Wiechert, *Schriften der Phys.-Okon. Ges. zu Konigsberg* 38 (1897), 3.
[15] W. Kaufmann, *Ann. der Phys. und Chem.* 61 (1897), 544.
[16] J. J. Thomson, *Phil. Mag.* 44 (1897), 310–12.
[17] ———, *Phil. Mag.* 48 (1899), 547.
[18] Ibid., p. 565.
[19] G. Jaffe, *J. Chem. Educ.* 29 (1952), 230.
[20] T. Carlyle, "On History," in *The Varieties of History*, ed. F. Stern (New York: Vintage, 1973).

What Is An Elementary Particle?

by STEVEN WEINBERG



WHEN A STRANGER, hearing that I am a physicist, asks me in what area of physics I work, I generally reply that I work on the theory of elementary particles. Giving this answer always makes me nervous. Suppose that the stranger should ask, "What is an elementary particle?" I would have to admit that no one really knows.

Let me declare first of all that there is no difficulty in saying what is meant by a *particle.* A particle is simply a physical system that has no continuous degrees of freedom except for its total momentum. For instance, we can give a complete description of an electron by specifying its momentum, as well as its spin around any given axis, a quantity that in quantum mechanics is discrete rather than continuous. On the other hand, a system consisting of a free electron and a free proton is not a particle, because to describe it one has to specify the momenta of both the electron and the proton—not just their sum. But a bound state of an electron and a proton, such as a hydrogen atom in its state of lowest energy, is a particle. Everyone would agree that a hydrogen atom is not an *elementary* particle, but it is not always so easy to make this distinction, or even to say what it means.

Research supported in part by the Robert A. Welch Foundation and NSF Grant PHY 9511632.

James Chadwick who discovered the neutron in 1932. (Courtesy AIP Meggers Gallery of Nobel Laureates)

FOR THE FIRST FEW decades of this century there did not seem to be any trouble in saying what is meant by an elementary particle. J. J. Thomson could use the electric field in a cathode-ray tube to pull electrons out of atoms, so atoms were not elementary. Nothing could be pulled or knocked out of electrons, so it seemed that electrons were elementary. When atomic nuclei were discovered in Ernest Rutherford's laboratory in 1911, it was assumed that they were not elementary, partly because it was known that some radioactive nuclei emit electrons and other particles, and also because nuclear charges and masses could be explained by assuming that nuclei are composed of two types of elementary particles: light, negatively charged electrons and heavy, positively charged protons.

Even without a definite idea of what is meant by an elementary particle, the idea that all matter consists of just two types of elementary particle was pervasive and resilient in a way that is difficult to understand today. For instance, when neutrons were discovered by James Chadwick in 1932, it was generally assumed that they were bound states of protons and electrons. In his paper announcing the discovery, Chadwick offered the opinion: "It is, of course, possible to suppose that the neutron is an elementary particle. This view has little to recommend it at present, except the possibility of explaining the statistics of such nuclei as N^{14}." (One might have thought this was a pretty good reason: molecular spectra had revealed that the N^{14} nucleus is a boson, which is not possible if it is a bound state of protons and electrons.) It was the 1936 discovery of the charge independence of nuclear forces by Merle Tuve et al. that showed clearly that neutrons and protons have to be treated in the same way; if protons are elementary, then neutrons must be elementary too. Today, in speaking of protons and neutrons, we often lump them together as *nucleons.*

This was just the beginning of a great increase in the roster of so-called elementary particles. Muons were added to the list in 1937 (though their nature was not understood until later), and pions and strange particles in the 1940s. Neutrinos had been proposed by Wolfgang Pauli in 1930, and made part of beta-decay theory by Enrico Fermi in 1933, but were not detected until the Reines-Cowan experiment of 1955. Then in the late 1950s the use of particle accelerators and bubble chambers revealed a great number of new particles, including mesons of spin higher than 0 and baryons of spin higher than 1/2, with various values for charge and strangeness.

On the principle that—even if there are more than two types of elementary particles—there really should not be a great number of types, theorists speculated that most of these particles are composites of a few types of elementary particles. But such bound states would have to be bound very deeply, quite unlike atoms or atomic nuclei. For instance, pions are much lighter than nucleons and antinucleons, so if the pion were a bound state of a nucleon and an antinucleon, as proposed by Fermi and Chen-Ning Yang, then its binding energy would have to be large enough to cancel almost all of the mass of its constituents. The composite nature of such a particle would be far from obvious.

How could one tell which of these particles is elementary and which composite? As soon as this question was asked, it was clear that the old answer—that particles are elementary if you can't knock anything out of them—was inadequate. Mesons come out when protons collide with each other, and protons and antiprotons come out when mesons collide with each other, so which is a composite of which? Geoffrey Chew and others in the 1950s turned this dilemma into a point of principle, known as "nuclear democracy," which held that every particle may be considered to be a bound state of any other particles that have the appropriate quantum numbers. This view was reflected decades later in a 1975 talk to the German Physical Society by Werner Heisenberg, who reminisced that:

> In the experiments of the fifties and sixties . . . many new particles were discovered with long and short lives, and no unambiguous answer could be given any longer to the question about what these particles consisted of, since this question no longer has a rational meaning. A proton, for example, could be made up of neutron and pion, or Lambda-hyperon and kaon, or out of two nucleons and an antinucleon; it would be simplest of all to say that a proton just consists of continuous matter, and all these statements are equally correct or equally false. The difference between elementary and composite particles has thus basically disappeared. And that is no doubt the most important experimental discovery of the last fifty years.


P. Ehrenfest, Jr.

Werner Heisenberg, left, talking with Neils Bohr at the Copenhagen Conference, Bohr Institute, 1934.
(Courtesy AIP Emilio Segrè Visual Archives)

LONG BEFORE Heisenberg reached this rather exaggerated conclusion, a different sort of definition of elementary particle had become widespread. From the perspective of quantum field theory, as developed by Heisenberg, Pauli, and others in the period 1926–34, the basic ingredients of Nature are not particles but fields; particles such as the electron and photon are bundles of energy of the electron and the electromagnetic fields. It is natural to define an elementary particle as one whose field appears in the fundamental field equations—or, as theorists usually formulate these theories, in the Lagrangian of the theory. It doesn't matter if the particle is heavy or light, stable or unstable—if its field appears in the Lagrangian, it is elementary; if not, not.

This is a fine definition if one knows the field equations or the Lagrangian, but for a long while physicists didn't. A fair amount of theoretical work in the 1950s and 1960s went into trying to find some objective way of telling whether a given particle type is elementary or composite when the underlying theory is

We will not be able to say which particles are elementary until we have a final theory of force and matter.

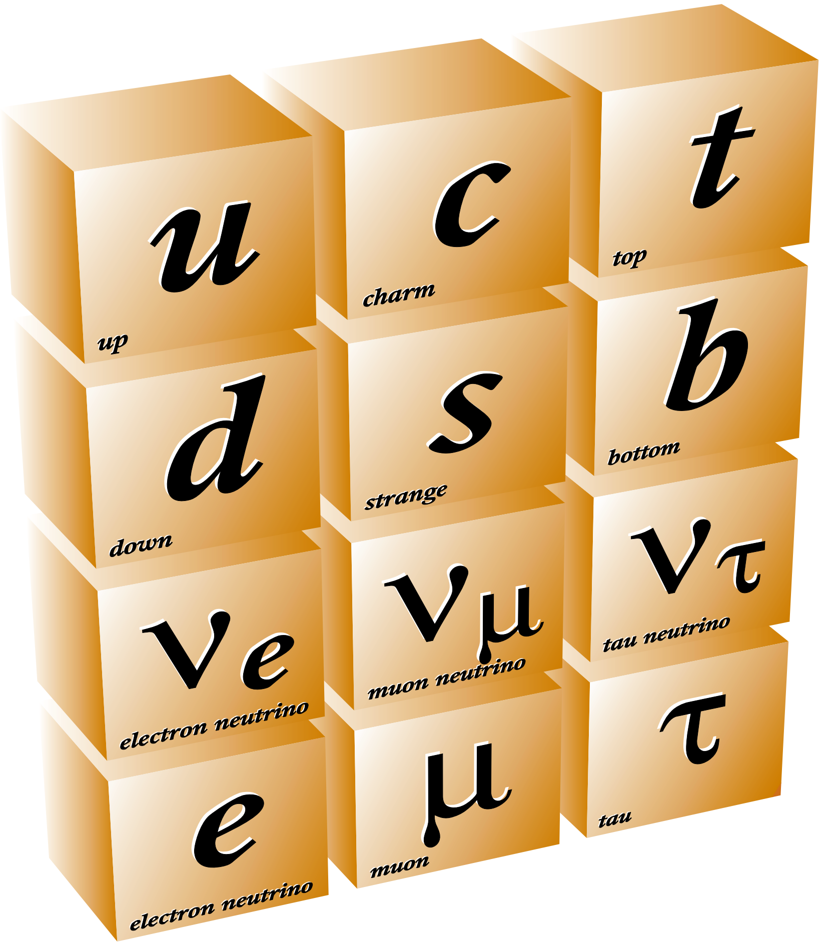
not known. This turned out to be possible in certain circumstances in nonrelativistic quantum mechanics, where an elementary particle might be defined as one whose coordinates appear in the Hamiltonian of the system. For instance, a theorem due to the mathematician Norman Levinson shows how to count the numbers of stable nonelementary particles minus the number of unstable elementary particles in terms of changes in phase shifts as the kinetic energy rises from zero to infinity. The trouble with using this theorem is that it involves the phase shifts at infinite energy, where the approximation of nonrelativistic potential scattering clearly breaks down.

I worried about this a good deal in the 1960s, but all I could come up with was a demonstration that the deuteron is a bound state of a proton and neutron. This was not exactly a thrilling achievement—everyone had always assumed that the deuteron is a bound state—but the demonstration had the virtue of relying only on nonrelativistic quantum mechanics and low-energy neutron-proton scattering data, without any specific assumptions about the Hamiltonian or about what happens at high energy. There is a classic formula that gives the spin triplet *s*-wave neutron-proton scattering length in terms of the nucleon mass and the deuteron binding energy, but the derivation of this formula actually relies on the assumption that the deuteron is a bound state. If we assume instead that the free-particle part of the Hamiltonian contains an elementary deuteron state, then this formula for the scattering length becomes incorrect, and instead we get a formula for the scattering length in terms of the nucleon mass, the deuteron binding energy, and the fraction of the time that the deuteron spends as an elementary particle (that is, the absolute value squared of the matrix element between the physical deuteron state and the elementary free-deuteron state). Comparing this formula with experiment showed that the deuteron spends most of its time as a composite particle. Unfortunately, arguments of this sort cannot be extended to deeply bound states, such as those encountered in elementary particle physics.

The lack of any purely empirical way of distinguishing composite and elementary particles does not mean that this distinction is not useful. In the 1970s the distinction between elementary and composite particles seemed to become much clearer, with the general acceptance of a quantum field theory of elementary particles known as the Standard Model. It describes quark, lepton, and gauge fields, so these are the elementary particles: six varieties or "flavors" of quarks, each coming in three colors; six flavors of leptons, including the electron; and twelve gauge bosons, including the photon, eight gluons, and the W^+, W^-, and Z^0 particles. The proton and neutron and all of the hundreds of mesons and baryons discovered after World War II are not elementary after all; they are composites of quarks and gluons, not because we can knock quarks and gluons out of them, which is believed to be impossible, but because that is the way they appear in the theory.

The one uncertain aspect of the Standard Model is the mechanism that breaks the electroweak gauge symmetry and gives the *W* and *Z* particles their masses, thereby adding an extra helicity state to what would have been the two helicities of a massless *W* or *Z* particle of spin 1. Theories of electroweak symmetry breakdown fall into two categories, according to whether these extra helicity states are elementary, as in the original form of the Standard Model, or composite, as in so-called technicolor theories. In a sense, the prime task driving the design of both the Large Hadron Collider and the ill-fated SSC was to settle the question of whether the extra helicity states of the *W* and *Z* particles are elementary or composite particles.

THIS MIGHT have been the end of the story, but since the late 1970s our understanding of quantum field theory has taken another turn. We have come to understand that particles may be described at sufficiently low energies by fields appearing in so-called effective quantum field theories, whether or not these particles are truly elementary. For instance, even though nucleon and pion fields do not appear in the Standard Model, we can calculate the rates for processes involving low-energy pions and nucleons by using an effective quantum field theory that involves pion and nucleon fields rather than quark and



Elementary particles today. There are three known families of quarks and leptons in the Standard Model.

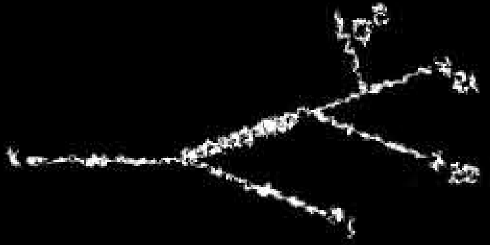
gluon fields. In this field theory pions and nucleons are elementary, though nuclei are not. When we use a field theory in this way, we are simply invoking the general principles of relativistic quantum theories, together with any relevant symmetries; we are not really making any assumption about the fundamental structures of physics.

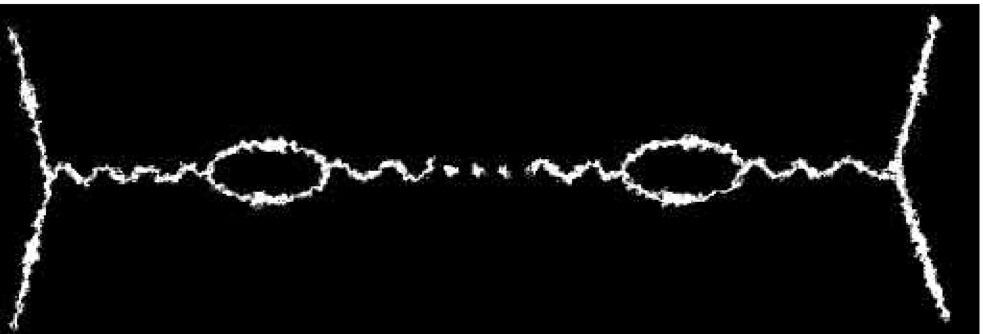
From this point of view, we are entitled only to say that the quarks and gluons are more elementary than nucleons and pions, because their fields appear in a theory, the Standard Model, that applies over a much wider range of energies than the effective field theory that describes nucleons and pions at low energy. We cannot reach any final conclusion about the elementarity of the quarks and gluons themselves. The Standard Model itself is probably only an effective quantum field theory, which serves as an approximation to some more fundamental theory whose details would be revealed at energies much higher than those available in modern accelerators, and which may not involve quark, lepton, or gauge fields at all.

One possibility is that the quarks and leptons and other particles of the Standard Model are themselves composites of more elementary particles. The fact that we see no structure in the quarks and leptons only tells us that the energies involved in their binding must be quite large—larger than several trillion electron volts. But so far no one has worked out a convincing theory of this sort.

We will not be able to give a final answer to the question of which particles are elementary until we have a final theory of force and matter. When we have such a theory, we may find that the elementary structures of physics are not particles at all. Many theorists think that the fundamental theory is something like a superstring theory, in which quarks, leptons, etc. are just different modes of vibration of the strings. It seems impossible in principle to identify one set of strings as truly elementary, because, as recently realized, different string theories with different types of strings are often equivalent.

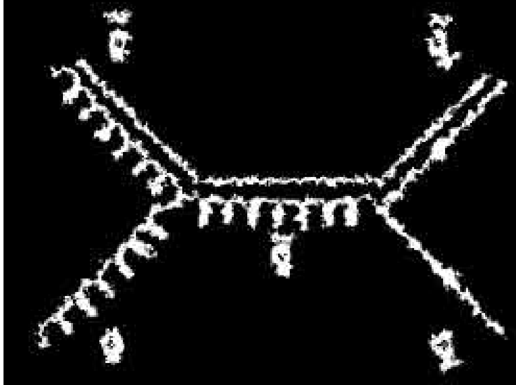
There is a lesson in all this. The task of physics is not to answer a set of fixed questions about Nature, such as deciding which particles are elementary. We do not know in advance what are the right questions to ask, and we often do not find out until we are close to an answer.



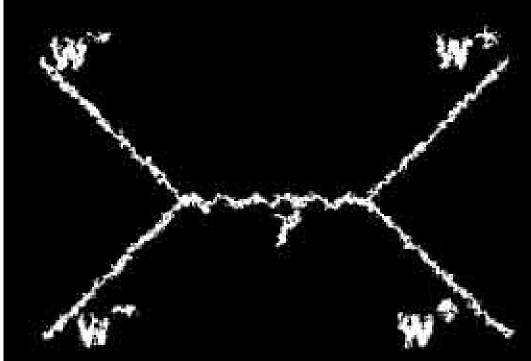


Elementary Particles: Yesterday, Today, and Tomorrow

by CHRIS QUIGG

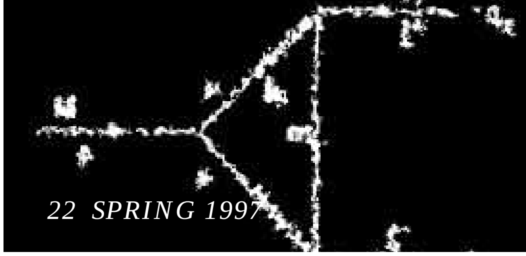


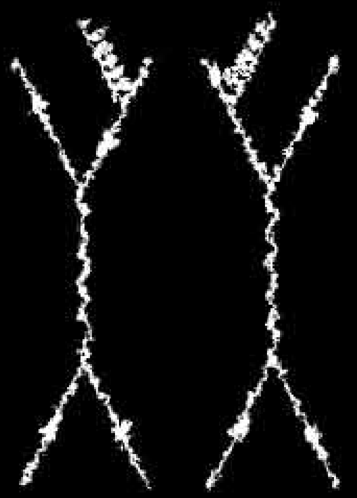
WITHIN THE LIFETIME of my grandparents, there lived distinguished scientists who did not believe in atoms. Within the lifetime of my children, there lived distinguished scientists who did not believe in quarks. Although we can trace the notion of fundamental constituents of matter—minimal parts—to the ancients, the experimental reality of the atom is a profoundly modern achievement. The experimental reality of the quark is more modern still.

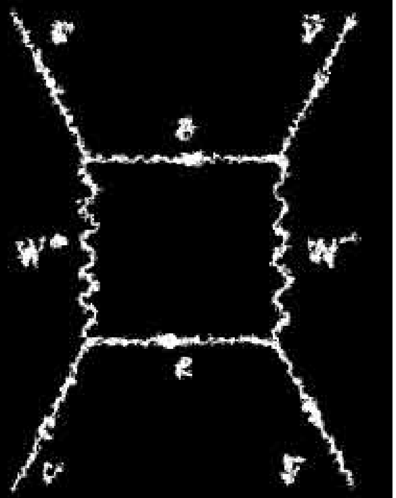


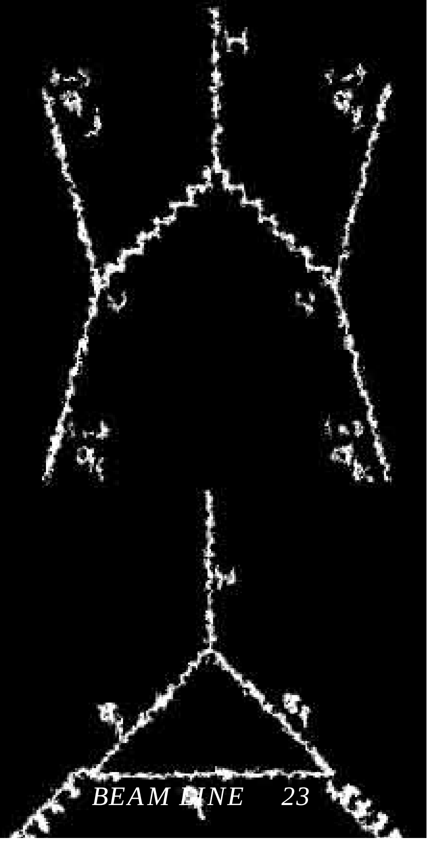
Through the end of the nineteenth century, controversy seethed over whether atoms were real material bodies or merely convenient computational fictions. The law of multiple proportions, the indivisibility of the elements, and the kinetic theory of gases supported the notion of real atoms, but it was possible to resist because no one had ever seen an atom. One of the founders of physical chemistry, Wilhelm Ostwald, wrote influential chemistry textbooks that made no use of atoms. The physicist, philosopher, and psychologist Ernst Mach likened “artificial and hypothetical atoms and molecules” to algebraic symbols, tokens devoid of physical reality that could be manipulated to answer questions about nature.

Atoms became irresistibly real when they began to come apart, with the discovery of the electron that we celebrate in









this special anniversary issue. In the end the atomists won not because they could see atoms—atoms are far too small to see—but because they learned to determine the size and weight of a single atom. In 1908 Jean-Baptiste Perrin established that the erratic “Brownian” movement of microscopic particles suspended in liquid was caused by collisions with molecules of the surrounding medium. This demonstration of the mechanical effects of tiny atoms and molecules effectively ended skepticism about their physical reality. Ostwald announced his conversion in 1909, the year he won the Nobel Prize. Mach went to his grave in 1916, still fighting a futile rear-guard action.

It is tempting to date the vanishing of resistance to the quark model to the discovery of the *J*/ψ particle in November 1974, but a look at the theoretical papers in the famous January 6, 1975, issue of *Physical Review Letters* will remind us that the epiphany wasn’t quite universal. The observation of the ψ',

Contents

Physical Review Letters VOLUME 34, NUMBER 1 6 JANUARY 1975

Atoms, Molecules, and Related Topics

Direct Observation of Atomic Energy Level Shifts in Two-Photon Absorption . . . . P. F. Liao and J. E. Bjorkholm 1

Plasmas, Fluids, Thermodynamics, and Related Topics

Steady-State Distributions of Interacting Discrete Vortices . . . . D. L. Book, Shalom Fisher, and B. E. McDonald 4
Spin Diffusion and Spin Echoes in Superfluid ^{3}He . . . . Roland Combescot 8

Solids

Propagating Order-Parameter Collective Modes in Superconducting Films . . . . R. V. Carlson and A. M. Goldman 11
Substrate Bias Effects on Electron Mobility in Silicon Inversion Layers at Low Temperatures . . . A. B. Fowler 15
X-Ray Inelastic Scattering of Li Metal . . . . P. Eisenberger, P. M. Platzman, and P. Schmidt 18
Sharp Optical Spectra of Impurities in Metals. . . . . R. A. Tilton and C. P. Flynn 20
Metal-Insulator Transition in Metal–Rare-Gas Alloys . . . . D. J. Phelps, R. Avci, and C. P. Flynn 23
Ferromagnetic Transition in V-Fe . . . . Helmut Claus 26

Nuclei

Exact-Finite-Range Coupled-Reaction-Channel Calculation Applied to the Analysis of Reaction ^{48}Ca(^{16}O, ^{15}C)^{49}Ti . . . . T. Udagawa, T. Tamura, and K. S. Low 30
Complete Isobaric Quintet. . . . . R. G. H. Robertson, W. S. Chien, and D. R. Goosman 33

Elementary Particles and Fields

Are the New Particles Baryon-Antibaryon Nuclei? . . . . Alfred S. Goldhaber and Maurice Goldhaber 36
Interpretation of a Narrow Resonance in e^+e^- Annihilation . . . . J. Schwinger 37
Possible Explanation of the New Resonance in e^+e^- Annihilation . . . . S. Borchardt, V. S. Mathur, and S. Okubo 38
Model with Three Charmed Quarks. . . . . R. Michael Barnett 41
Heavy Quarks and e^+e^- Annihilation . . . . Thomas Appelquist and H. David Politzer 43
Is Bound Charm Found? . . . . A. De Rújula and S. L. Glashow 46
Possible Interactions of the *J* Particle . . . . H. T. Nieh, Tai Tsun Wu, and Chen Ning Yang 49
Remarks on the New Resonances at 3.1 and 3.7 GeV . . . . C. G. Callan, R. L. Kingsley, S. B. Treiman, F. Wilczek, and A. Zee 52

Comments

Intermediate Boson in the Fermion-Current Model of Neutral Currents . . . . J. J. Sakurai 56

Errata

Microwave Dimensional Resonances in Large Electron-Hole Drops in Germanium . . . . R. S. Markiewicz, J. P. Wolfe, and C. D. Jeffries 59

Proceedings of the XIIIth International
Conference on High-Energy Physics,
Berkeley, California
August 31–September 7, 1966

Contents

Foreword
Preface
Conference Organizing Comitttee
Introductory Session
  CURRENT TOPICS IN PARTICLE PHYSICS
  M. Gell-Mann, Speaker
Session 1
  FUNDAMENTAL THEORETICAL QUESTIONS
  M. Froissart, Rapporteur
Session 2
  WEAK INTERACTIONS
  N. Cabibbo, Rapporteur
Session 3
  CURRENT COMMUTATORS
  R. F. Dashen, Rapporteur
Session 4
  EXPERIMENTS IN T, C VIOLATION
  V. L. Fitch, Rapporteur
Session 5
  OTHER ASPECTS OF WEAK INTERACTIONS: HIGH ENERGY NEUTRINO PHYSICS AND QUESTIONS OF C, P, T NONINVARIANCE
  T. D. Lee, Rapporteur
Session 6
  ELECTRODYNAMIC INTERACTIONS
  S. Drell, Rapporteur
Session 7
  BOSON RESONANCES
  G. Goldhaber, Rapporteur
Session 8
  PERIPHERAL COLLISIONS AT INTERMEDIATE ENERGIES
  J. D. Jackson, Rapporteur
Session 9a
  S = 0 BARYON RESONANCES
  P. G. Murphy, Rapporteur
Session 9b
  BARYON RESONANCES WITH S ≠ 0
  M. Ferro-Luzzi, Rapporteur
Session 10
  SYMMETRIES AND THE STRONG INTERACTIONS
  R. H. Dalitz, Rapporteur
Session 11
  DYNAMICS OF STRONG INTERACTIONS
  F. Low, Rapporteur
Session 12
  HADRON COLLISIONS AT VERY HIGH ENERGIES
  L. Van Hove, Rapporteur

a second new particle that was obviously related to the *J*/*ψ*, made the notion of quarks as mechanical objects irresistible to all but an obdurate few. The holdouts were either converted or consigned to a just irrelevance by the discovery of charm eighteen months later.

MEETING THE QUARK

My first contact with quarks came during the summer of 1966, as I was about to begin graduate school in Berkeley. Before I had set foot in a classroom, the Thirteenth International Conference on High Energy Physics took place on campus, a gathering of about four hundred scientists from around the world. Though attendance was by invitation, with strict national quotas, I could present myself at the front door of Wheeler Auditorium in the morning and obtain a day pass that allowed me to sit inconspicuously in the back of the room and watch the proceedings. Except for what I had learned that summer working through two little books by Richard Feynman, I knew nothing of the interactions between particles, or even what the particles were like. So there I was, *tabula rasa* among the experts.

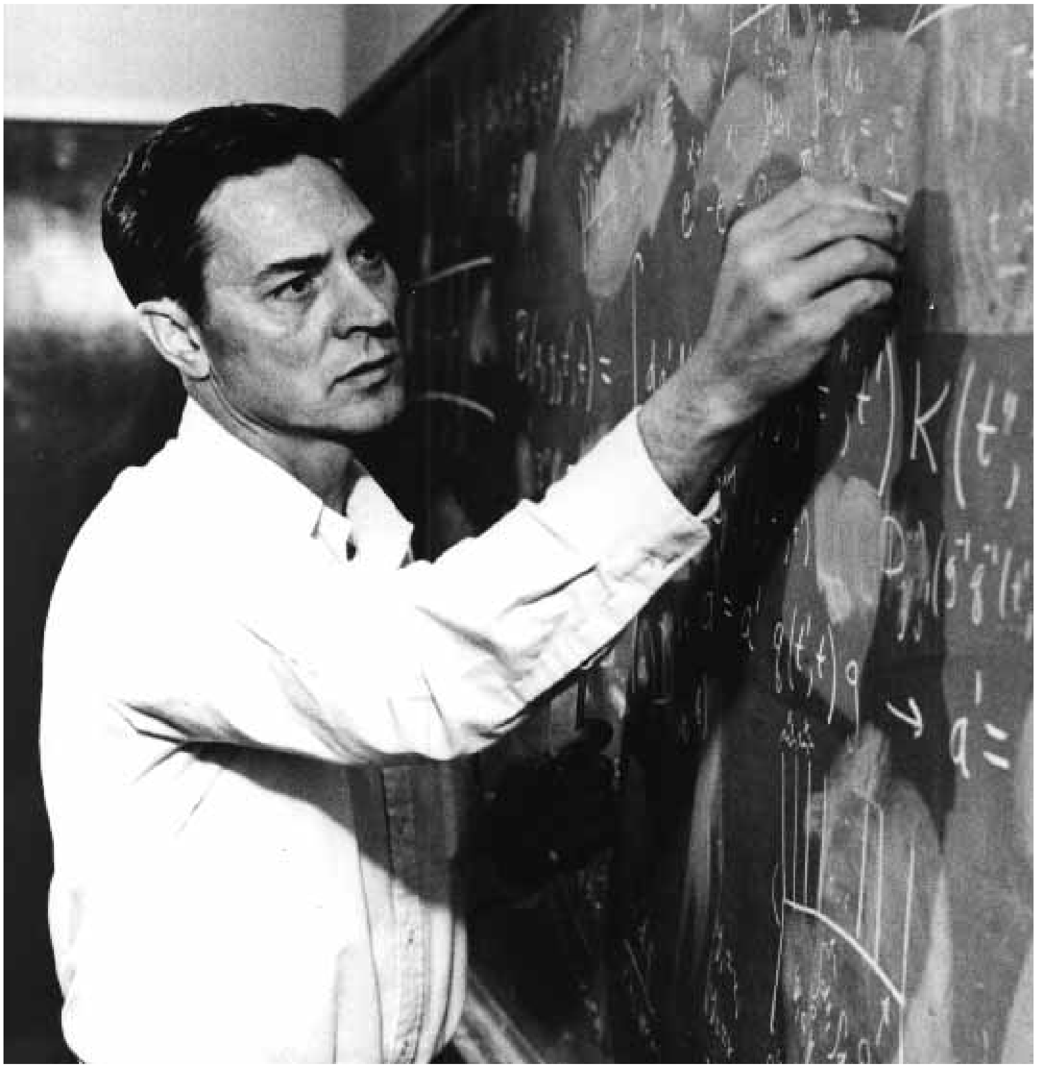
I could understand a little of the opening address by Murray Gell-Mann and a talk on symmetries by Richard Dalitz of Oxford. Both of them talked—rather cautiously, it seemed—about hypothetical objects called quarks as fundamental constituents of the proton and neutron and all the other strongly interacting particles. Although the idea that three quarks made up a proton while a quark and antiquark made up a meson brought order to a lot of information, it was clear that nobody had any idea how this could happen and whether there could be a self-consistent theory. And besides, no one had seen a quark.

Richard Dalitz in 1961.

Courtesy G.-C. Wick and AIP Emilio Segrè Archives

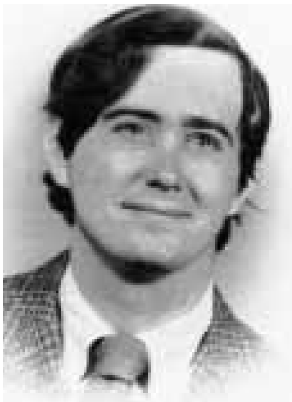
Just as the Greek atomists had their opponent in Anaxagoras, who advocated an infinite progression of

Opposite: Physicists attending the 1966 International Conference on High Energy Physics at Berkeley, California. *(Courtesy Lawrence Berkeley National Laboratory)*



Right: Geoffrey Chew in the 1960s. *(Courtesy Lawrence Berkeley National Laboratory and AIP Emilio Segrè Visual Archives)*

seeds within seeds—and no minimal parts—the quark advocates had to face the challenge of "nuclear democracy." Berkeley, it turned out, was the hotbed of an opposing point of view: that there were no fundamental constituents, that all the composite "elementary" particles were somehow made out of each other in an intricate interplay called the bootstrap. Gell-Mann deflected this challenge by repeatedly stressing that quarks didn't have to be real to be useful and that if the mesons and baryons were made up of "mathematical quarks," then the quark model might perfectly well be compatible with the bootstrap hypothesis.



Courtesy C. Quigg

The author in 1970, as a fresh Ph.D. and research associate in the Institute for Theoretical Physics at the State University of New York, Stony Brook.

There was also the question of how to deal with interactions, with theorists divided into sects promoting "*S*-matrix theory," or "field theory," or "Lagrangian field theory," or "abstract field theory." Gell-Mann urged the partisans to stop wasting their breath on sectarian quarrels and to pool their energies to campaign for a higher-energy accelerator that would enable us to really learn more about the basic structure of matter. That accelerator sweeps across the prairie outside my office window.

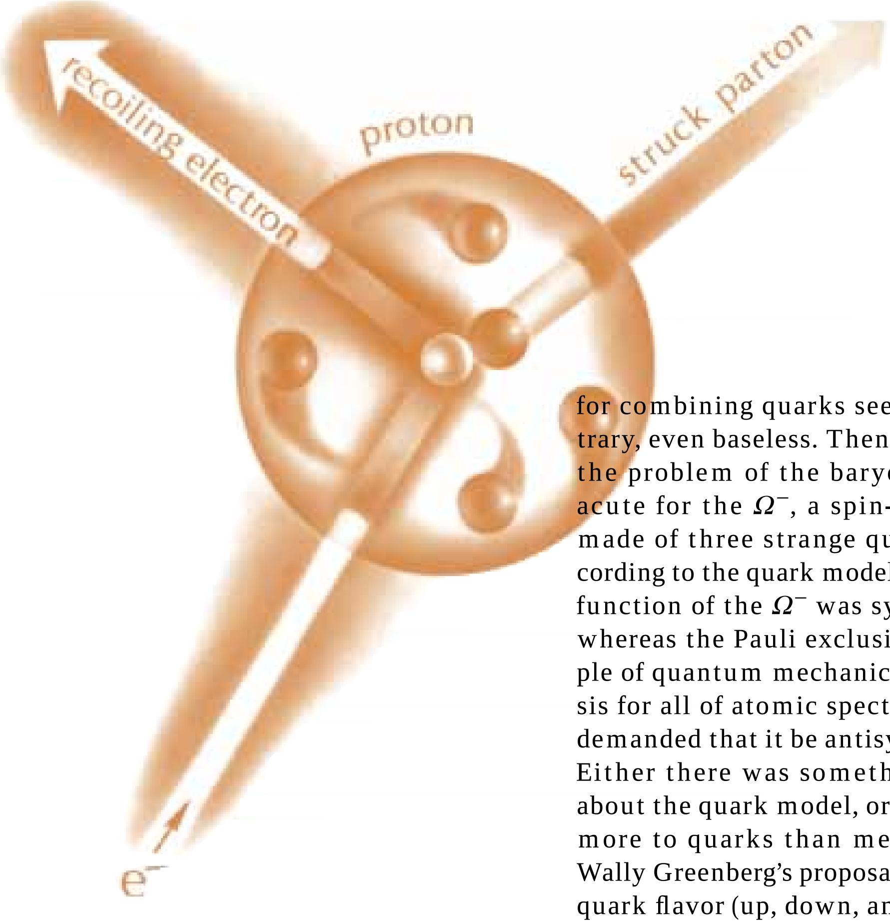
QUARKS IN BERKELEY?

Berkeley was indeed the Mother Church of the *S*-matrix bootstrap denomination. I don't think quarks were ever mentioned in Geoff Chew's course on the dynamics of strong interactions. Even in Dave Jackson's more catholic version of the course, quarks appeared only once, on a list of term-paper topics at the end of the year. But that was only part of the story. Learning about other approaches to particles and interactions was not only encouraged, it was obligatory. Berkeley graduate students were expected to follow two year-long courses in field theory. The Rad Lab was a center of hadron spectroscopy where the quark model was discussed as a classification tool. In the spring of 1968, George Zweig flew up from Caltech every Friday to teach a graduate seminar on the quark model. George was one of the inventors of quarks. He also knew everything about resonance spectra and decays, and he gleefully showed us how much a simple quark model could explain.

Courtesy G. Zweig

George Zweig in 1965.

What the quark model couldn't explain was itself: "How could this be true?" was the question everyone had to ask. Until the interactions of quarks could be understood, the rules



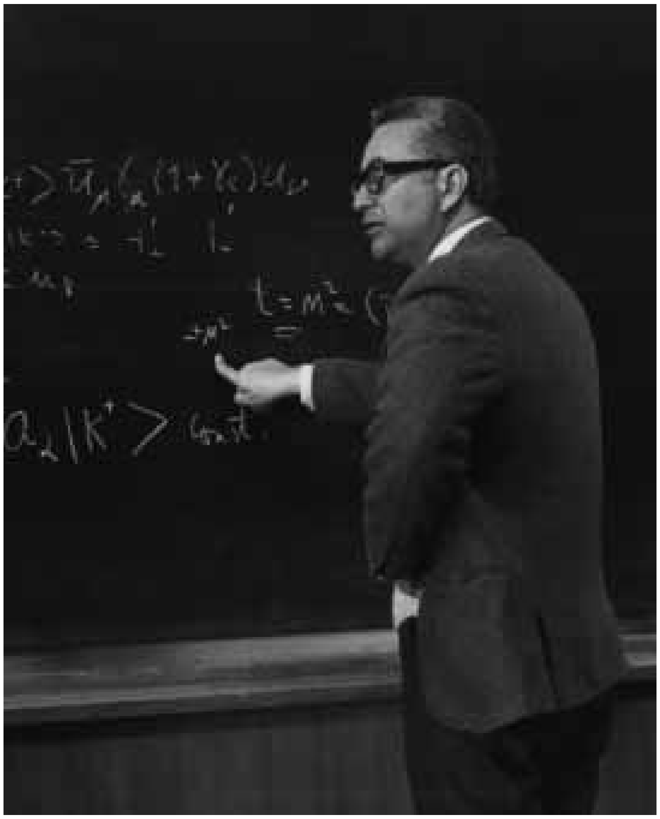
for combining quarks seemed arbitrary, even baseless. Then there was the problem of the baryons, most acute for the Ω^-, a spin-$\frac{3}{2}$ particle made of three strange quarks. According to the quark model, the wave function of the Ω^- was symmetric, whereas the Pauli exclusion principle of quantum mechanics—the basis for all of atomic spectroscopy—demanded that it be antisymmetric. Either there was something dicey about the quark model, or there was more to quarks than met the eye. Wally Greenberg's proposal that each quark flavor (up, down, and strange) came in three distinguishable "colors," and that antisymmetry in color brought the quark model into conformance with the exclusion principle, seemed to many like invoking the tooth fairy. But in one of those delicious ironies that make research so interesting, when we learned to measure the number of colors of each quark species, it really was three. And color would turn out to be the key to explaining how the quark model could be true.

Richard Feynman lecturing on his parton model at SLAC in October 1968.

Henry W. Kendall

The other evidence that drew attention to quarks arose from the MIT-SLAC experiments in which Jerry Friedman, Henry Kendall, Dick Taylor, and their colleagues studied the structure of the proton. To the prepared mind, the high rate of inelastic collisions they observed showed that there were within the proton tiny charged bodies. No mind was more prepared to take the leap than Feynman's. Feynman presented his interpretation at a SLAC colloquium that occasioned my first pilgrimage across the Bay. The colloquium was then held in the evening after what has been described to me as a vintner's dinner. Whatever the reason, I remember both speaker and audience as extremely exuberant. If an electron scattered from one of the hypothetical tiny charged bodies, not the whole proton, it was easy to understand why the inelastic cross section was so large. Instead of measuring the delicacy of the proton, the MIT and SLAC experimenters were measuring the hardness of the little bits. Feynman wasn't prepared to say what the tiny charged parts of the proton were, so he called them "partons." Everyone in the room must have thought, "Quarks?"

Before long, Bj Bjorken and Manny Paschos had worked out the consequences of the quark-parton model for electron scattering and neutrino scattering. The success of their predictions added to a gathering crisis. If the quark-partons acted as if they



Murray Gell-Mann in 1972. (Courtesy CERN)

were free, independent objects when examined by energetic electrons, why didn't the quarks come out and show themselves? Gell-Mann derided Feynman's picture as the "put-on" model. Many theorists of my generation found great sport in showing that Bjorken's scaling law, which was implied by the parton model, wasn't possible in this or that interacting field theory. Like the quark model of the hadron resonances, the parton model could explain many things, but it couldn't explain itself.

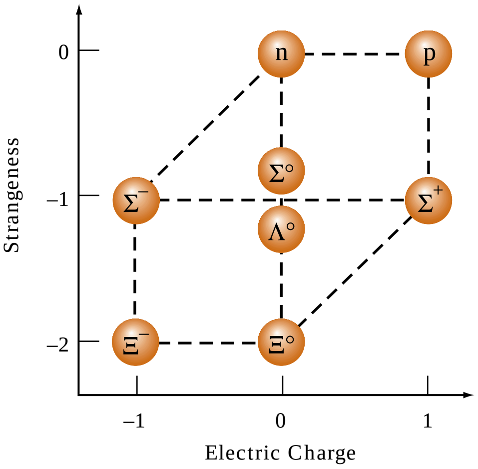
DYNAMICS, DYNAMICS, DYNAMICS!

Some of the reasons why it took so long for the idea of quarks to be accepted have to do with the human frailties of obtuseness, or obstinacy, or preoccupation with other matters. But others, the reasons of real importance, reflect the standards of scientific evidence. The repeated failure to find any free quarks sustained the idea that quarks were computational fictions. The main sticking-point was the absence of any understanding of how quarks behave as free and independent objects in hard collisions, and yet form composites in which they are permanently confined. Without an understanding of dynamics, quarks were a story, not a theory.

The great illumination came in 1973, when David Gross and Frank Wilczek in Princeton and David Politzer at Harvard found that, alone among field theories, non-Abelian gauge theories could reconcile the permanent confinement of quarks with the relative independence the parton model presumes. In these theories the interaction between two quarks diminishes when they are close together, but becomes an ineluctable pull when the quarks move apart. This "asymptotic freedom" of the strong interaction is just what was needed to understand the MIT-SLAC results—not just in a useful cartoon, but in a real theory.

In what seemed like the blink of an eye, a new theory of the strong interactions was codified. Gell-Mann named it quantum chromodynamics (QCD) to celebrate the central role of color as the strong-interaction charge and perhaps to express the hope that it would become as fertile and robust as quantum electrodynamics, the phenomenally successful theory of electrons and photons. Soon precise predictions emerged for the subtle deviations from Bjorken scaling that QCD predicted.

Even before the scaling violations implied by QCD were established through painstaking experimental effort, asymptotically free gauge theories gave us license to take the quark model and the parton picture seriously. All at once, what we had gingerly called "as-if" models took on new meaning. Now, the J/ψ was such a thunderbolt that it needed no theoretical stage-dressing to help it set the community of particle physicists on its ear. Yet it was the insight of asymptotic freedom that prepared us to read the clues charmonium offered, and change forever the way we think about the structure of matter.



The first single-electron event from Gargamelle. The electron's trajectory goes from left to right, beginning at the arrow's tip. The haloed black circles are lights to illuminate the bubble-chamber liquid. (Courtesy CERN)

QUARKS, LEPTONS, GAUGE FIELDS

Today's elementary particles, the leptons (ν_e, e), (ν_μ, μ), (ν_τ, τ), and the quarks (u,d), (c,s), (t,b), form one of the pillars of our understanding of matter and energy. To the limits of our resolution, they are all spin-$\frac{1}{2}$ particles with no internal structure. The quarks are color triplets that experience the strong interactions. The leptons, which have no color charge, do not.

The top quark has so far been seen in such small numbers that we haven't yet examined it as closely as the others. If top is as ephemeral as we think, with a lifetime less than a trillionth of a trillionth of a second, it is the purest quark—the only one that does not form mesons or baryons. We know a great deal about the tau neutrino from the study of τ and Z decays, but it would still be satisfying to execute a "three-neutrino experiment," in which a beam of tau neutrinos interacts with a target to produce tau leptons that live for a millimeter or two before they decay. The DONUT (Direct Observation of NU-Tau) experiment being commissioned at Fermilab should observe about 150 examples of the $\nu_\tau \rightarrow \tau$ transition.

The other essential foundation for our current understanding is the notion that symmetries—gauge symmetries—determine the character of the fundamental interactions. Like QCD, the electroweak theory fashioned by Sheldon Glashow, Steven Weinberg, and Abdus Salam is a non-Abelian gauge theory. The electroweak theory got its own boost in the summer of 1973 when André Lagarrigue and his colleagues in the Gargamelle bubble-chamber experiment at CERN announced the first observation of weak neutral-current interactions. Although it would take the discovery of the weak-force particles W and Z and many years of study, culminating in the contributions of the Z factories at CERN and SLAC, to show how successful a creation the electroweak theory is, it was clear very soon that the gauge-field-theory approach to the interactions of quarks and leptons was the right path.

The electroweak theory supplies a clue of profound significance: our world must have both quarks and leptons. Unless each pair of leptons (like the electron and its neutrino) is accompanied by a pair of quarks (like up and down), quantum corrections will clash with the symmetries from which the electroweak theory is derived, leaving it inconsistent. I take this constraint as powerful encouragement for a family relationship joining quarks and leptons, and for a unified theory of the strong, weak, and electromagnetic interactions.

Have we found all the quarks and leptons? We do not really know. Precision measurements of the width of the Z resonance assure us that there are no more normal generations with very light neutrinos. But there could well be new families of quarks and leptons in which all the members are too massive to be produced in Z decays. We don't know yet whether the neutrinos have any mass. If they do, we need to learn whether each neutrino is its own antiparticle.

Even if we have already met all the quarks and leptons, we have good reason to be open to the possibility of new kinds of matter. The astrophysical case for dark matter in the galaxy is persuasive, and the evidence that most of the matter in the Universe is both nonluminous and unlike the stuff we know is highly suggestive. Supersymmetry, which is for the moment the most popular candidate to extend the electroweak theory, implies a greatly expanded list of elementary particles, including spin-zero partners of the quarks and leptons.

> *The idea that elementary constituents of matter interact according to the dictates of gauge symmetries has become the organizing principle of particle physics, as important to our field as evolution is to biology.*

If we take as our goal not merely describing the world as we find it, but understanding why the Universe is the way it is, the array of elementary particles presents us with many challenges. What makes a top quark a top quark, or an electron an electron? Can we calculate the masses of the quarks and leptons and the relative strengths of the weak transitions between quark flavors? Why are there three generations?

ANOTHER LAYER OF STRUCTURE?

No experimental evidence except the history of molecules and atoms and protons suggests that quarks and leptons are composite. However, there is an undeniable aesthetic allure to the notion that a complex world may arise out of the combinatoria of a few simple parts. If today's elementary particles are composite, we might be able to compute their masses, understand the trebling of generations, and decipher the relationship of quarks to leptons.

Some specific currents in theoretical research also lead toward composite quarks and leptons. In dynamical theories of electroweak symmetry breaking such as technicolor, the longitudinal components of the weak gauge bosons are composite. Why not the quarks and leptons, too? And a new approach to supersymmetric model-building, in which strong gauge interactions break the supersymmetry, suggests that some of the quarks may be composite.

Composite models of quarks and leptons must differ in a crucial way from familiar dynamical pictures. In QCD the pions are the lightest—nearly massless—particles, while the proton mass is set by the scale of binding energy. A theory of quark and lepton compositeness must deliver fermions much lighter than the (several TeV, at least) binding energy of the constituents. Without a specific composite model, we have no theoretical clue for the scale on which we might resolve structure in our elementary particles. Nevertheless, we can characterize the experimental signatures that composite quarks and leptons would leave.

At energies approaching such a compositeness scale, quarks and leptons that have size will interact at such short distances that they interpenetrate and rearrange, or even exchange, their constituents. In quark-quark scattering, the conventional gluon exchange of QCD would be supplemented by a contact interaction whose strength is determined by the size of the quarks. In $\bar{p}p$ collisions, this new contribution would lead to an excess of hadron jets at large values of the transverse energy, where quark-antiquark scattering is the dominant reaction. Typically, the angular distribution of the jets will differ from the shape QCD predicts. If quarks and leptons have common constituents, a similar excess will be seen in dilepton production from the elementary process $\bar{q}q \rightarrow \ell^+\ell^-$. At still higher energies, we would expect to see the effects of excited quarks and leptons. Finally, at energies well above this compositeness scale, quarks and leptons would begin to manifest form factors characteristic of their size.

Since I first met the quark, charm, beauty, top, and the rest have become my friends and teachers—in fact, they have taken over my life. The idea that elementary constituents of matter interact according to the dictates of gauge symmetries has become the organizing principle of particle physics, as important to our field as evolution is to biology. I don't know how far the revolution of quarks and leptons and gauge bosons will carry us, but I have a wish for the decade ahead: that we will learn the true nature of electroweak symmetry breaking and begin to understand the individuality of the quarks and leptons. And I hope we will look back with pleasure and satisfaction at how passionate and optimistic—and naïve—we were in 1997.

THE INDUSTRIAL STRENGTH

by MICHAEL RIORDAN

MORE THAN A DECADE before J. J. Thomson discovered the electron, Thomas Edison stumbled across a curious effect, patented it, and quickly forgot about it. Testing various carbon filaments for electric light bulbs in 1883, he noticed a tiny current trickling in a single direction across a partially evacuated tube into which he had inserted a metal plate. Two decades later, British entrepreneur John Ambrose Fleming applied this effect to invent the “oscillation valve,” or vacuum diode—a two-terminal device that converts alternating current into direct. In the early 1900s such rectifiers served as critical elements in radio receivers, converting radio waves into the direct current signals needed to drive earphones.

In 1906 the American inventor Lee de Forest happened to insert another electrode into one of these valves. To his delight, he discovered he could influence the current flowing through this contraption by changing the voltage on this third electrode. The first vacuum-tube amplifier, it served initially as an improved rectifier. De Forest promptly dubbed his triode the audion and applied for a patent. Much of the rest of his life would be spent in forming a series of shaky companies to exploit this invention—and in an endless series of legal disputes over the rights to its use.

These pioneers of electronics understood only vaguely—if at all—that individual subatomic particles were streaming through their devices. For them, electricity was still the fluid (or fluids) that the classical electrodynamicists of the nineteenth century thought to be related to stresses and disturbances in the luminiferous æther. Edison, Fleming and de Forest might have been dimly aware of Thomson’s discovery, especially after he won the 1906 Nobel

Copyright © 1996 by Michael Riordan. Adapted in part from *Crystal Fire: The Birth of the Information Age,* by Michel Riordan and Lillian Hoddeson, to be published in 1997 by W. W. Norton & Co.

PARTICLE

Prize in physics. But this knowledge had yet to percolate out of academic research labs such as the Cavendish and into industrial workshops. Although he had earned a Ph.D. in physics from Yale, in his daily practice de Forest remained pretty much a systematic tinkerer in the Edisonian vein, trying endless variations on his gadgets in his halting attempts to improve their performance.

VOLUMES COULD be written about the practical applications that owe their existence to the understanding of electricity as a stream of subatomic particles rather than a continuous fluid. While the telephone clearly antedated the discovery of the electron, for example, its modern manifestations—cellular and touchtone phones, telefax machines, satellite communications—would be utterly impossible without such knowledge. And the ubiquitous television set is of course just a highly refined version of the cathode-ray tube that Thomson used to determine the charge-to-mass ratio of his beloved corpuscle. The field of electronics, a major subfield of electrical engineering today, grew up in the twentieth century around this new conception of electricity, eventually taking its name in the 1920s from the particle at its core. (We are perhaps fortunate that Thomson did not prevail in his choice of nomenclature!)

In parallel with the upsurge of electronics, and in some part due to it, came a sweeping transformation of industrial research in America. Once the main province of highly individualistic inventors searching for a fruitful breakthrough, technology development slowly became an organized practice performed by multidisciplinary teams of salaried scientists and engineers working in well-equipped industrial labs. As the century waxed and quantum mechanics emerged to explain the mysterious behavior of electrons, atoms and molecules, these researchers increasingly sported advanced degrees in physics or chemistry. A deeper understanding of the scientific principles governing the behavior of matter gradually became indispensable to the practice of industrial research. As the noted historian of technology Thomas Hughes put it, “Independent inventors had manipulated machines and dynamos; industrial scientists would manipulate electrons and molecules.”

Few examples illustrate this evolutionary transformation better than the case of the vacuum-tube amplifier. For almost a decade after de Forest invented it, his audion found little use beyond low-voltage applications in wireless receivers—as a detector of weak radio signals. He simply did not understand that the gas remaining in his tube was impeding the flow of electrons from filament to plate. At the higher voltages required for serious amplification, say in telephone communications, the device began, as one observer noted, “to fill with blue haze, seem to choke, and then transmit no further speech until the incoming current had been greatly reduced.”

One corporation extremely interested in amplifying telephone signals was the American Telephone and Telegraph Company, then seeking to develop a suitable “repeater” for transcontinental phone service.

J. J. Thomson inspecting electron tubes in 1923 with Frank Jewett, the first president of Bell Labs. (Courtesy AT&T Archives and AIP Niels Bohr Library)

Among its leading scientists was Frank Jewett, then working in the engineering department of its Western Electric Division. In 1902 he had earned a Ph.D. in physics from the University of Chicago, doing his research under Albert Michelson and befriending Robert Millikan. Harboring a hunch that the electrical discharges in evacuated tubes might serve as the basis for a suitable repeater, Jewett approached his old chum, who in 1911 sent one of his brightest graduate students, Harold Arnold, to Western Electric. Here was a young man steeped in the new thinking, who had just spent several years measuring the charges of individual electrons on oil droplets.

When de Forest demonstrated his audion to Western Electric scientists and engineers in October 1912, Arnold was present. He diagnosed the blue haze as due to the recombination of gas molecules that had been ionized by energetic electrons. Then he solved its problems by use of high vacuum, an oxide-coated filament, and other modifications dictated by a superior understanding of the electronic discharge. (A similar development occurred simultaneuosly at General Electric, but it lost the ensuing patent fight to AT&T, which had wisely purchased the appropriate rights to de Forest's patents.)

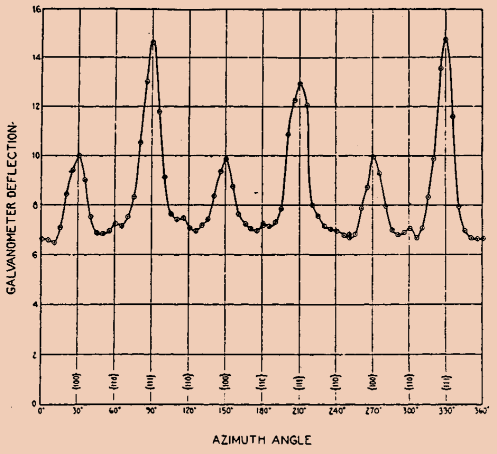
Within a year Western Electric was making "high-vacuum thermionic tubes" that served as active elements in excellent telephone repeaters. At the grand opening of the Panama-Pacific Expositon held in San Francisco on January 15, 1915, Alexander Graham Bell inaugurated the nation's first coast-to-coast telephone service, talking to his former assistant Thomas Watson in New York. Recalling this event in his autobiography, Millikan observed that "the electron—up to that time largely the plaything of the scientist—had clearly entered the field as a patent agent in the supplying of man's commercial and industrial needs."

Thus convinced of the value of scientific research in an industrial setting, Western Electric incorporated its engineering department as a separate entity—the Bell Telephone Laboratories—in 1925, naming Jewett its first president. The very next year, as an outgrowth of their research on the performance of vacuum tubes (also called electron tubes), Clinton Davisson and Lester Germer established the wave nature of electrons, which had been predicted a few years earlier by Louis de Broglie. For his pivotal work on electron diffraction, Davisson was to share the 1937 Nobel Prize in physics with the British scientist George Thomson, son of J. J.

Quantum mechanics soon explained the behavior not only of electrons in atoms but of the large ensembles of them that swarm about freely within metals. Based on the theoretical work of Enrico Fermi and Paul Dirac, Bell Labs physicists eventually figured out why an oxide-coating worked so well on tungsten filaments of vacuum tubes. It helped to lower the work function of the

Right: Clinton Davisson and Lester Germer with the apparatus they used to establish the wave nature of electrons.
(Courtesy AT&T Archives)

Bottom: Graph from their 1927 Nature *article showing diffraction peaks observed in electron scattering from a nickel crystal.*



metal, thereby making it easier for electrons to escape from the surface—and substantially reducing the amount of power needed to heat a filament. Such a fundamental understanding of the physics of electrons proved crucial to further engineering advances in vacuum tubes that saved AT&T millions of dollars annually.

IN THE LATE 1920S and early 1930s, Felix Bloch, Rudolph Peierls, Alan Wilson and other European physicists laid the foundations of modern solid-state physics in their theoretical studies of how waves of electrons slosh about within the periodic potentials encountered inside crystalline materials. Their work resulted in a theory of solids in which there are specific allowed (or forbidden) energy levels—called "bands"—that electrons can (or cannot) occupy, analogous to the Bohr orbitals of early quantum theory. Combined with practical methods of calculating these band structures in actual substances, pioneered by Eugene Wigner, band theory fostered a better understanding of why certain materials act as electrical conductors and others as insulators. And, in a decade when electron tubes reigned supreme as the active components of electronic circuits, band theory began to elucidate the properties of intermediate materials called semiconductors, whose myriad spawn would eventually supplant these tubes throughout electronics.

World War II spurred tremendous practical advances in the technology of semiconductors, largely due to the fact that microwave receivers needed rectifiers able to operate above a few hundred megahertz, where electron tubes had proved useless. Crystal rectifiers, with a delicate metal point pressed into a germanium or silicon surface, filled the gap nicely. By the end of the War, methods of purifying and doping these substances to make easily controlled, well-understood semiconductors had been perfected by scientists at such secret enclaves as the Rad Lab at MIT and Britain's Telecommunications Research Establishment at Great Malvern.

No laggard itself in these pursuits, Bell Labs led the way during the postwar years in applying wartime insights and technologies to the creation of practical new semiconductor components. "The quantum physics approach to structure of matter has brought about greatly increased understanding of solid-state phenomena," wrote its vice president Mervin Kelly—another of Millikan's grad students—in 1945, authorizing formation of a solid-state physics group. "The modern conception of the constitution of solids that has resulted indicates that there are great possibilities of producing new and useful properties by finding physical and chemical methods of controlling the arrangement of the atoms and electrons which compose solids."

The most important postwar breakthrough to occur at Bell Labs was the invention of the transistor in late 1947 and early 1948 by John Bardeen, Walter Brattain, and William Shockley. And a key to their interpretation of transistor action was a new physical phenomenon Shockley dubbed "minority carrier injection"—in which electrons and positively charged quantum-mechanical entities called "holes" can flow by diffusion in the presence of one another. Once again, a detailed scientific understanding of how individual subatomic particles (and what, in certain respects, act like their antiparticles) behave proved crucial to a pivotal advance in electronics.

The transistor happened along at a critical juncture in technological history. For the electronic digital computers that also emerged from wartime research could not have evolved much further without it. The thousands of bulky, fragile electron tubes used in such early,

Lee de Forest, inventor of the vacuum-tube amplifier, and Bell Labs President Mervin Kelly. (Courtesy AT&T Archives)

room-filling computers as the ENIAC and UNIVAC burned out with all-too-frustrating frequency. Only large corporations, the armed services and government agencies could afford these massive, power-hungry monstrosities and the vigilant staff to keep them operating. "It seems to me," Shockley conjectured in December 1949, "that in these robot brains the transistor is the ideal nerve cell."

But the transistor has proved to be much more than merely a replacement for electron tubes and electromechanical switches. Shrunk to less than a ten-thousandth of its original size and swarming by the millions across the surfaces of microchips, it has opened up entirely unexpected realms of electronic possibility, which even the most farsighted could not have anticipated during those booming postwar years. The transistor was, as historians Ernest Braun and Stuart MacDonald observed, "the harbinger of an entirely new sort of electronics with the capacity not just to influence an industry or a scientific discipline, but to change a culture."

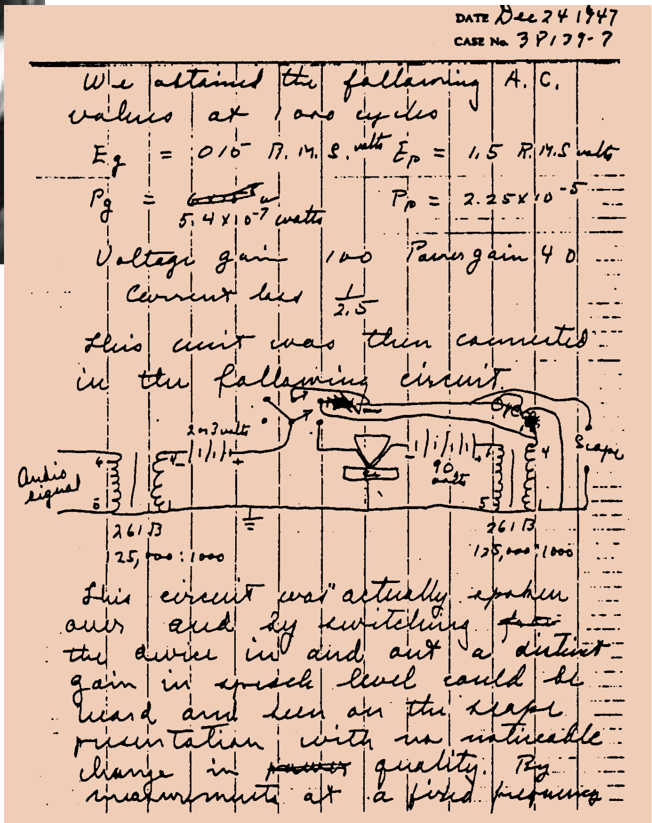
Characteristically, particle physicists were among the first to glimpse the potential ramifications of this revolutionary new solid-state amplifier. "I would be very anxious to do some experimenting to learn about the techniques of your new Germanium triods," wrote Fermi to Shockley in early January 1949 (misspelling the final word). After receiving a few samples and testing them at his University of Chicago laboratory, Fermi replied, "They really are very fine gadgets, and I hope very much that they might be useful in our work."

THOMSON'S DISCOVERY triggered a spectacular century of innovation in both science and technology. Paced by increasingly detailed knowledge of the electron's properties and behavior, scientists and engineers developed many other advanced devices—lasers, light-emitting diodes, microwave tubes, solar cells and high-speed microprocessors, to name several—that are essential to modern computing and global communications. Today we know the mass and charge of the electron to better than seven significant figures. Aided by quantum mechanics, we can accurately calculate its energy levels in all kinds of atoms, molecules and solid-state substances. Largely taken for granted, such information is crucial for the precision control of electrons at the submicron scales that characterize many leading-edge technologies.

Of critical importance in attaining this deep understanding was the ease with which electrons can be detached from other forms of matter and manipulated using electromagnetic fields. Such features were readily apparent in Thomson's landmark experiments, for example, and Millikan exploited them in his research. In certain key instances the energy required corresponds to that of photons in visible light. This unique partnership between the electron and photon (whose centennial we will also celebrate in the not-too-distant future) is central to much of

Left: John Bardeen, William Shockley, and Walter Brattain, who shared the 1956 Nobel Prize in physics for the invention of the transistor. (Courtesy AT&T Archives)

Bottom: Page from Brattain's lab notebook that records the 23 December 1947 demonstration of the first transistor. (Courtesy AT&T Archives)



DATE Dec 24 1947
CASE No. 38139-7

We obtained the following A.C. values at 1000 cycles

E_g = .015 R.M.S. volts E_p = 1.5 R.M.S volts

P_g = 5.4×10^{-7} watts P_p = 2.25×10^{-5}

Voltage gain 100 Power gain 40

Current loss $\frac{1}{2.5}$

This unit was then connected in the following circuit.

This circuit was actually spoken over and by switching the device in and out a distinct gain in speech level could be heard and seen on the scope presentation with no noticeable change in quality. By measurements at a fixed frequency

advanced technology. Perhaps this complementary relationship is one reason why, during the past decade or so, we have witnessed the emergence of a "photonics" industry that threatens to supplant electronics in some commercial sectors.

No comparable partnership exists, for example, between quarks and gluons. Quarks are *not* detachable, at least not yet, and gluon fields do not extend to infinity like the electromagnetic. It makes all the difference in the world. We still have only rough values for the quark masses, especially the up and down quarks that make up the bulk of ordinary matter. Only a few wild-eyed speculators dare to mention the possibility of "quarkonics" or "gluonics" industries. Maybe one day soon we will manipulate quarks for practical benefit, say by using high-energy electron beams. Perhaps, but I seriously doubt it will ever amount to much of an industry. For now, quarks and gluons remain the playthings of pure physics.

Just a hundred years after its discovery, the electron sits at the core of modern life. Where previous generations of writers, for example, used chemical inks and mechanical devices to cast their ideas onto paper (and deliver them to readers), I have composed this article while staring at a luminous screen—behind which a blazing stream of electrons traces out my fumbling thoughts as my fingertips tap keys that activate transistor-laden microprocessors deep within my computer. And some of you now read my words on luminous screens of your own, conveyed to your desks by surging rivers of electrons and photons pulsating as ones and zeroes through an intricate network that stretches into almost every corner of the globe. The only paper involved in the exchange is the pages marked with red ink that now lie crumpled in my wastebasket.

We all owe a debt to J. J. Thomson—and the scientists and engineers who followed the path that he pioneered—for taming the first subatomic particle and adapting its unique properties for the practical applications that are relentlessly redefining what it means to be human.

The Evolution of Particle Accelerators & Colliders

by WOLFGANG K. H. PANOFSKY

WHEN J. J. THOMSON discovered the electron, he did not call the instrument he was using an accelerator, but an accelerator it certainly was. He accelerated particles between two electrodes to which he had applied a difference in electric potential. He manipulated the resulting beam with electric and magnetic fields to determine the charge-to-mass ratio of cathode rays. Thomson achieved his discovery by studying the properties of the beam itself—not its impact on a target or another beam, as we do today. Accelerators have since become indispensable in the quest to understand Nature at smaller and smaller scales. And although they are much bigger and far more complex, they still operate on much the same physical principles as Thomson's device.

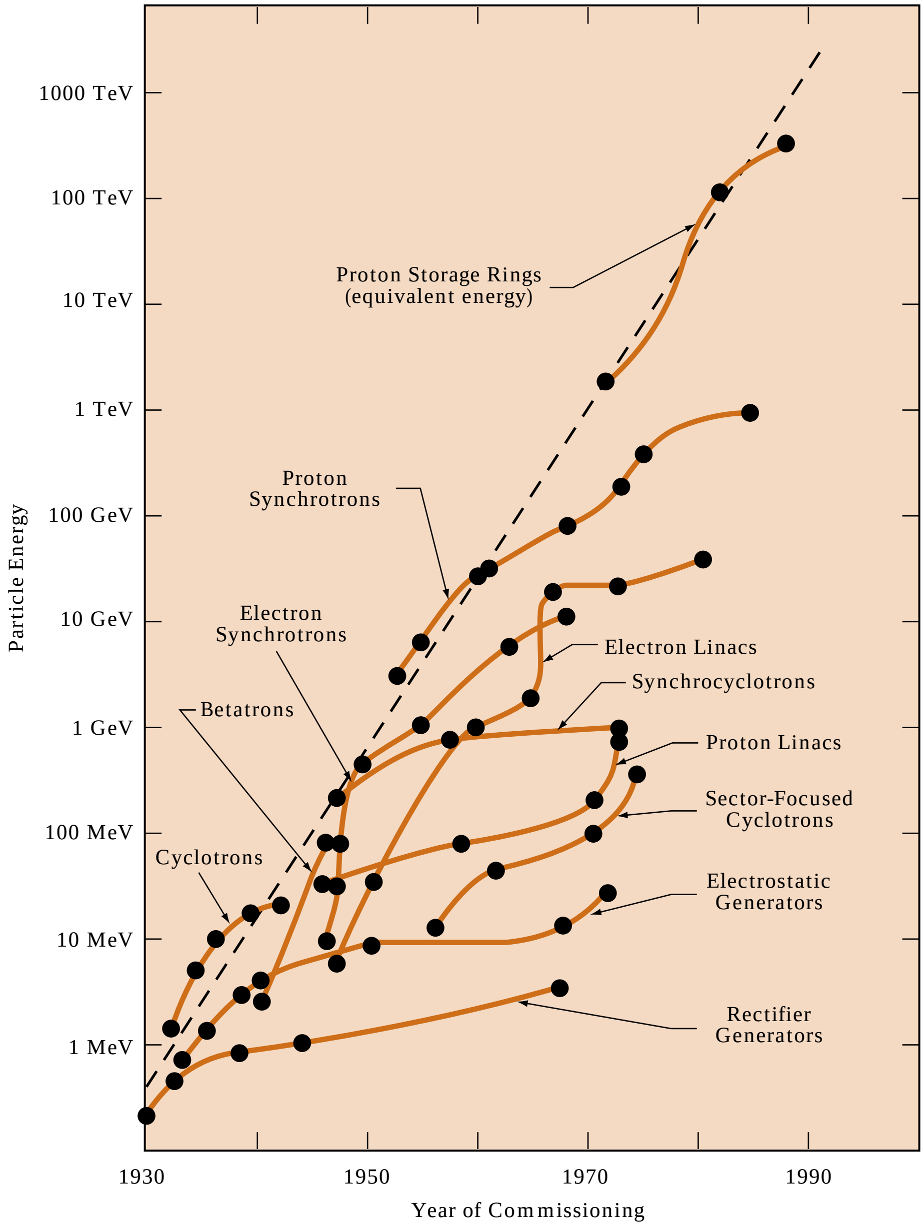
It took another half century, however, before accelerators became entrenched as the key tools in the search for subatomic particles. Before that, experiments were largely based on natural radioactive sources and cosmic rays. Ernest Rutherford and his colleagues established the existence of the atomic nucleus—as well as of protons and neutrons—using radioactive sources. The positron, muon, charged pions and kaons were discovered in cosmic rays.

One might argue that the second subatomic particle discovered at an accelerator was the neutral pion, but even here the story is more complex. That it existed had already been surmised from the existence of charged pions, and the occurrence of gamma rays in cosmic rays gave preliminary evidence for such a particle. But it was an accelerator-based experiment that truly nailed down the existence of this elusive object.

There followed almost two decades of accelerator-based discoveries of other subatomic particles originally thought to be elementary, notably the antiproton and the vector mesons. Most of these particles have since turned out to be composites of quarks. After 1970 colliders—machines using two accelerator beams in collision—entered the picture. Since then most, but certainly not all, new revelations in particle physics have come from these colliders.

IN CONSIDERING the evolution of accelerator and collider technology, we usually think first of the available energy such tools provide. Fundamentally, this is the way it should be. When the study of the atomic nucleus stood at the forefront of "particle physics" research, sufficient energy was needed to allow two nuclei—which are positively charged and therefore repel one another—to be brought close enough to interact. Today, when the components of these nuclei are the main objects of study, the reasons for high energy are more subtle. Under the laws of quantum mechanics, particles can be described both by their physical trajectory as well as through an associated wave whose behavior gives the probability that a particle can be localized at a given point in space and time. If the wavelength of a probing particle is short, matter can be examined at extremely small distances; if long, then the scale of things that can be investigated will be coarser. Quantum mechanics relates this wavelength to the energy (or, more precisely, the momentum) of the colliding particles: the greater the energy, the shorter the wavelength.

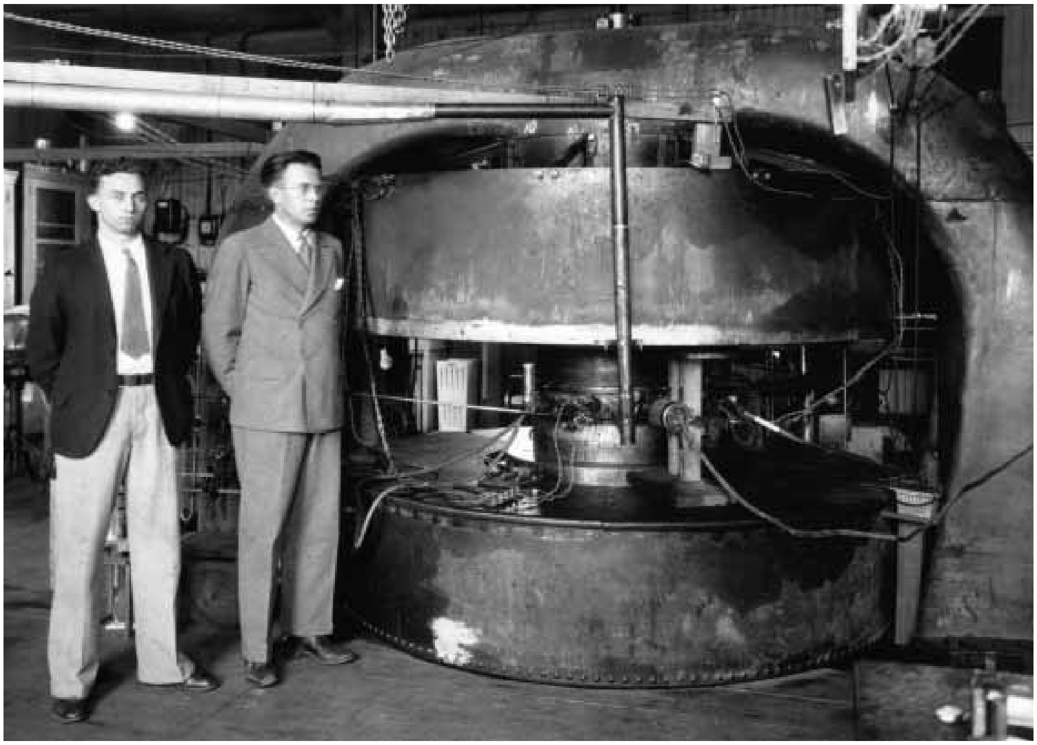
A "Livingston plot" showing accelerator energy versus time, updated to include machines that came on line during the 1980s. The filled circles indicate new or upgraded accelerators of each type.



This relationship can be expressed quantitatively. To examine matter at the scale of an atom (about 10^{-8} centimeter), the energies required are in the range of a thousand electron volts. (An electron volt is the energy unit customarily used by particle physicists; it is the energy a particle acquires when it is accelerated across a potential difference of one volt.) At the scale of the nucleus, energies in the million electron volt—or MeV—range are needed. To examine the fine structure of the basic constituents of matter requires energies generally exceeding a billion electron volts, or 1 GeV.

But there is another reason for using high energy. Most of the objects of interest to the elementary particle physicist today do not exist as free particles in Nature; they have to be created artificially in the laboratory. The famous $E = mc^2$ relationship governs the collision energy E required to produce a particle of mass m. Many of the most interesting particles are so heavy that collision energies of many GeV are needed to create them. In fact, the key to understanding the origins of many parameters, including the masses of the known particles, required to make today's theories consistent is believed to reside in the attainment of collision energies in the trillion electron volt, or TeV, range.

Our progress in attaining ever higher collision energy has indeed been impressive. The graph on the left, originally produced by M. Stanley Livingston in 1954, shows how the laboratory energy of the particle beams produced by accelerators has increased. This plot has been updated by adding modern developments. One of the first things to notice is that the energy of man-made accelerators has been growing exponentially in time. Starting from the 1930s, the energy has increased—roughly speaking—by about a factor of 10 every six to eight years. A second conclusion is that this spectacular achievement has resulted



M. Stanley Livingston and Ernest O. Lawrence, with their 27-inch cyclotron at Berkeley Radiation Laboratory. (Courtesy Lawrence Berkeley National Laboratory)

from a succession of technologies rather than from construction of bigger and better machines of a given type. When any one technology ran out of steam, a successor technology usually took over.

In another respect, however, the Livingston plot is misleading. It suggests that energy is the primary, if not the only, parameter that defines the discovery potential of an accelerator or collider. Energy is indeed required if physicists wish to cross a new threshold of discovery, provided that this threshold is defined by the energy needed to induce a new phenomenon. But there are several other parameters that are important for an accelerator to achieve—for example, the intensity of the beam, or the number of particles accelerated per second.

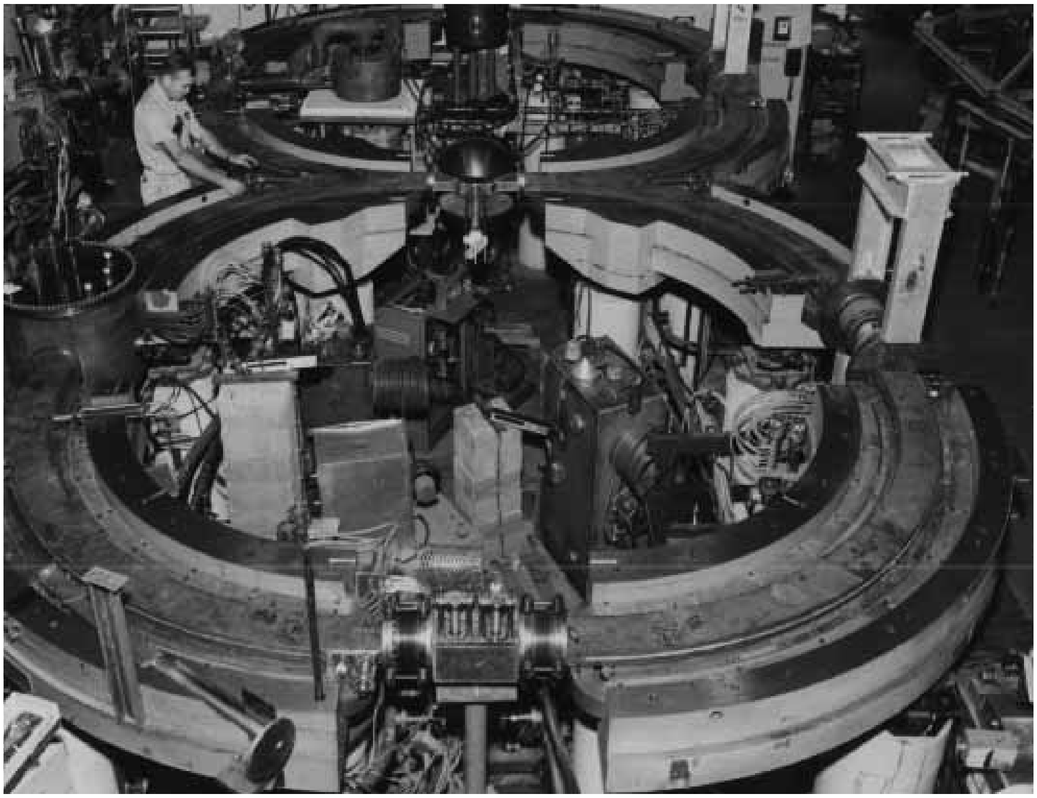
When the beam strikes a target, its particles collide with those in the target. The likelihood of producing a reaction is described by a number called the cross section, which is the effective area a target particle presents to an incident particle for that reaction to occur. The overall interaction rate is then the product of the beam intensity, the density of target particles, the cross section of the reaction under investigation, and the length of target material the incident particle penetrates. This rate, and therefore the beam intensity, is extremely important if physicists are to collect data that have sufficient statistical accuracy to draw meaningful conclusions.

Another important parameter is what we call the duty cycle—the percentage of time the beam is actually on. Unlike Thomson's device, most modern accelerators do not provide a steady flow of particles, generally because that would require too much electric power; instead, the beam is pulsed on and off. When physicists try to identify what reaction has taken place, one piece of evidence is whether the different particles emerge from a collision at the same time. Thus electronic circuits register the instant when a particle traverses a detector. But if the accelerator's duty cycle is small, then all the particles will burst forth during a short time interval. Therefore a relatively large number of accidental coincidences in time will occur, caused by particles emerging from different individual reactions, instead of from real coincidences due to particles emerging from a single event. If time coincidence is an important signature, a short duty cycle is a disadvantage.

Then there is the problem of backgrounds. In addition to the reaction under study, detectors will register two kinds of undesirable events. Some backgrounds arise from particles generated by processes other than the beam's interaction with the target or another beam—such as with residual gas, from "halo" particles traveling along the main beam, or even from cosmic rays. Other backgrounds stem from reactions that are already well understood and contain no new information. Accelerators differ in terms of the presence or absence of both kinds of backgrounds; their discovery potential differs accordingly. The amount and kinds of background are directly related to the ease of data analysis, the type of detector to be built, or whether the desired results can be extracted at all.

In general, as the energy increases, the number of possible reactions also increases. So does the burden on the discriminating power of detectors and on the data-analysis potential of computers that can isolate the "wheat" from the "chaff." With the growth in energy indicated by the Livingston plot, there had to be a parallel growth in the analyzing potential of the equipment required to identify events of interest—as well as a growth in the number of people involved in its construction and operation.

And finally there is the matter of economy. Even if a planned accelerator is technically capable of providing the needed energy, intensity, duty cycle, and low background, it still must be affordable and operable. The resources required—money, land, electric power—must be sufficiently moderate that the expected results will have



The first colliding-beam machine, a double-ring electron-electron collider, built by a small group of Princeton and Stanford physicists. (Courtesy Stanford University)

commensurate value. Of course “value” has to be broadly interpreted in terms not only of foreseeable or conjectured economic benefits but also of cultural values related to the increase in basic understanding. In view of all these considerations, the choice of the next logical step in accelerator construction is always a complex and frequently a controversial issue. Energy is but one of many parameters to be considered, and the value of the project has to be sufficiently great before a decision to go ahead can be acceptable to the community at large.

All these comments may appear fairly obvious, but they are frequently forgotten. Inventions that advance just one of the parameters—in particular, energy—are often proposed sincerely. But unless the other parameters can be improved at the same time, to generate an overall efficient complex, increasing the energy alone usually cannot lead to fundamentally new insights.

THE ENERGY that really matters in doing elementary particle physics is the collision energy—that is, the energy available to induce a reaction, including the creation of new particles. This collision energy is *less* than the laboratory energy of the particles in a beam if that beam strikes a stationary target. When one particle hits another at rest, part of the available energy must go toward the kinetic energy of the system remaining after the collision. If a proton of low energy E strikes another proton at rest, for example, the collision energy is $E/2$ and the remaining $E/2$ is the kinetic energy with which the protons move ahead. At very high energies the situation is complicated by relativity. If a particle of total energy E hits another particle of mass M, then the collision energy is given by $E_{coll} \sim (2Mc^2E)^{1/2}$, which is much less than $E/2$ for E much larger than Mc^2.

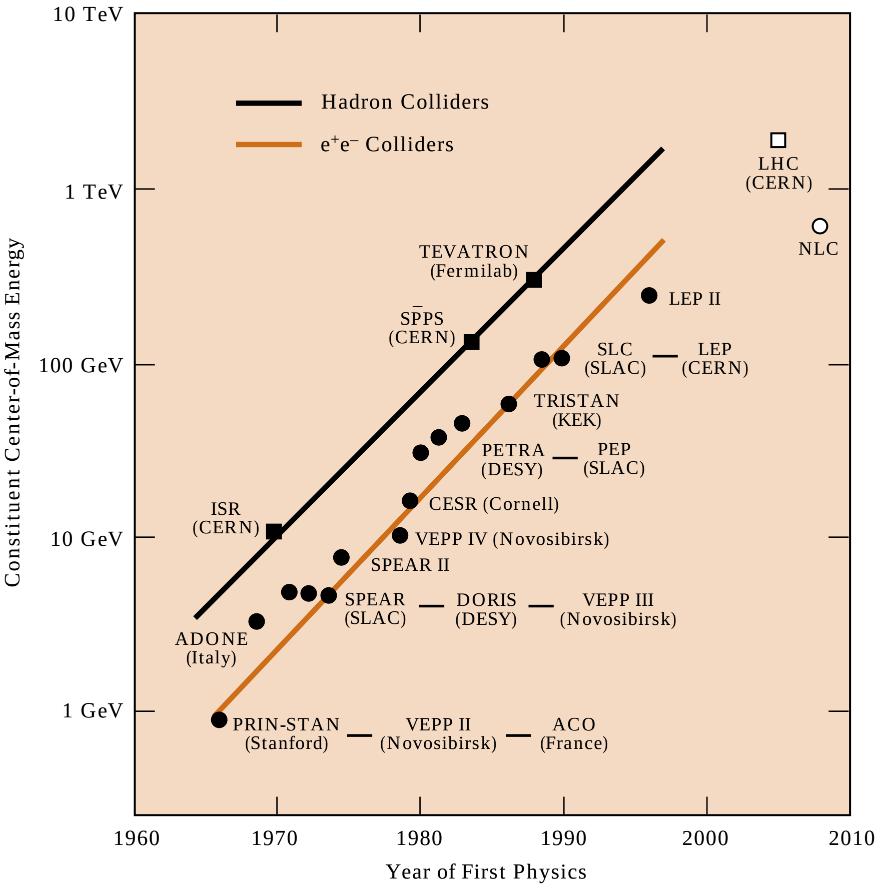
If two particles of equal mass traveling in opposite directions collide head on, however, the total kinetic energy of the combined system after collision is zero, and therefore the entire energy of the two particles becomes available as collision energy. This is the basic energy advantage offered by colliding-beam machines, or colliders.

The idea of colliding-beam machines is very old. The earliest reference to their possibility stems from a Russian publication of the 1920s; it would not be surprising if the same idea occurred independently to many people. The first collider actually used for particle-physics experiments, built at Stanford in the late 1950s, produced electron-electron collisions (see photograph on the left). Other early machines, generating electron-positron collisions, were built in Italy, Siberia and France. Since then there has been a plethora of electron-positron, proton-proton and proton-antiproton colliders.

There is another problem, however. If the particles participating in a collision are themselves composite—that is, composed of constituents—then the available energy must be shared among these constituents. The threshold for new phenomena is generally defined by the collision energy in the constituent frame: the energy that becomes available in the interaction between two individual constituents. Here there are major differences that depend on whether the accelerated particles are protons, deuterons, electrons or something else. Protons are composed of three quarks and surrounded by various gluons. Electrons and muons, as well as quarks and gluons, are considered pointlike, at least down to distances of 10^{-16} centimeter. Recognizing these differences, we can translate the Livingston plot into another chart (top right, next page) showing energy in the constituent frame versus year of operation for colliding-beam machines.

But the idea of generating higher collision energy via colliding beams

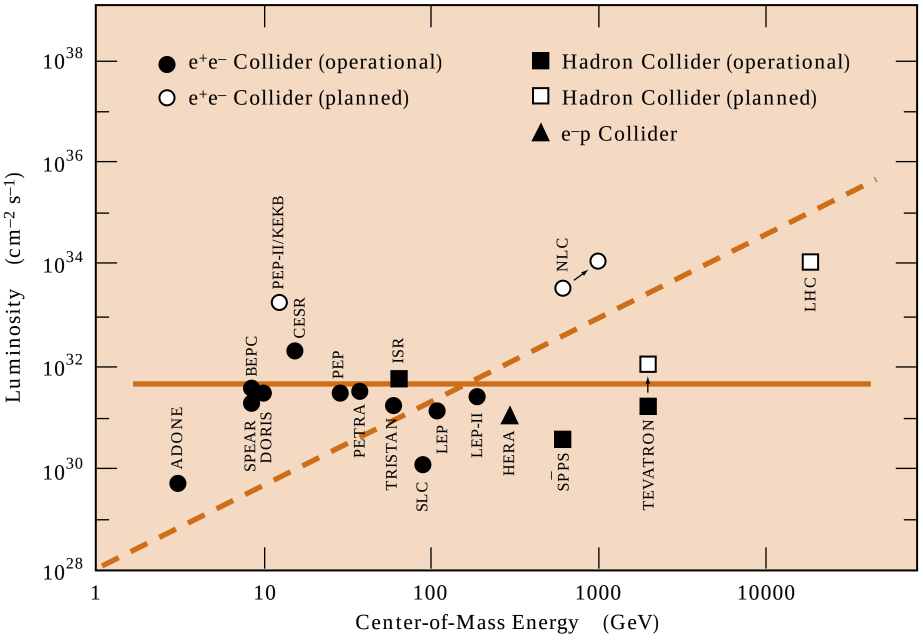
Right: The energy in the constituent frame of electron-positron and hadron colliders constructed (filled circles and squares) or planned. The energy of hadron colliders has here been derated by factors of 6–10 in accordance with the fact that the incident proton energy is shared among its quark and gluon constituents.



is worthless unless (as discussed above) higher interaction rates can be generated, too. To succeed, the density of the two beams must be high enough—approaching that of atoms in ordinary matter—and their interaction cross sections must be sufficient to generate an adequate data rate. In colliding-beam machines the critical figure is the luminosity *L*, which is the interaction rate per second per unit cross section. The bottom graph on this page illustrates the luminosity of some of these machines. In contrast to the constituent collision energy, which has continued the tradition of exponential growth begun in the Livingston plot, the luminosity has grown much more slowly. There are good reasons for this trend that I will discuss shortly.

Naturally there are differences that must be evaluated when choosing which particles to use in accelerators and colliders. In addition to the energy advantage mentioned for electrons, there are other factors. As protons experience the strong interaction, their use is desirable, at least in respect to hadron-hadron interactions. Moreover, the cross sections involved in hadron interactions are generally much larger than those encountered in electron machines, which therefore require higher luminosity to be equally productive.

Proton accelerators are generally much more efficient than electron machines when used to produce secondary beams of neutrons, pions, kaons, muons, and neutrinos. But electrons produce secondary beams that are sharply concentrated in the forward direction, and these beams are less contaminated by neutrons.



Peak luminosities achieved at existing colliders and values projected for planned or upgraded machines. The dashed line indicates luminosity increasing as the square of the center-of-mass energy Note that the rated machine energy has been used in calculating the abscissa. (Data updated courtesy Greg Loew, SLAC)

Far left: William Hansen (right) and colleagues with a section of his first linear electron accelerator, which operated at Stanford University in 1947. Eventually 3.6 m long, it could accelerate electrons to an energy of 6 MeV. (Courtesy Stanford University)

Left: Ernest Lawrence's first successful cyclotron, built in 1930. It was 13 cm in diameter and accelerated protons to 80 keV. (Courtesy Lawrence Berkeley National Laboratory)

When discussing the relative merits of electron and proton colliders, the background situation is complex because the factors that cause them are quite different. When accelerated, and especially when their path is bent by magnets, electrons radiate X rays in the form of synchrotron radiation. Protons usually have more serious interactions with residual gas atoms, and those that deviate from the nominal collider orbit are more apt to produce unwanted backgrounds from such causes.

A much more difficult—and to some extent controversial—subject is the comparison of the complexities of events initiated by electrons with those induced by hadrons in general, and protons in particular. Today particle physicists are usually, but not always, interested in the results of "hard" collisions between the elementary constituents (by which I mean entities considered to be pointlike at the smallest observable distances). Because protons are composite objects, a single hard collision between their basic constituents will be accompanied by a much larger number of extraneous "soft" collisions than is the case for electrons. Thus the fraction of interesting events produced in an electron machine is generally much larger than it is for proton machines. So the analysis load in isolating the "needle" from the "haystack" tends to be considerably more severe at hadron machines.

Beyond this problem is the matter of interpretability. When we use electrons to bombard hadron targets, be they stationary or contained in an opposing beam, we are exploring a complex structure with an (as-yet) elementary object whose behavior is well understood. Thus the information about the structure of the proton resulting from electron-proton collisions, for example, tends to be easier to interpret than the results from proton-proton collisions. All the above observations are generalities, of course, and there are numerous and important exceptions. For instance, if neutrinos or muons—copiously produced as secondary beams from proton machines—are used to explore the structure of hadrons, the results are complementary to those produced by electron beams.

Everything I have said about electrons is also true of muons. The use of muon beams offers significant advantages and disadvantages relative to electrons. The two lightest charged leptons, the electron and muon, experience essentially the same interactions. But muons, being heavier, radiate far less electromagnetic energy than do electrons of equal energy; therefore backgrounds from radiative effects are much lower. On the other hand, muons have a short lifetime (about 2 microseconds), whereas electrons are stable. Colliding-beam devices using muons must be designed to be compatible with this fact. In addition, the remnants of the muons that decay during acceleration and storage constitute a severe background. Thus, while the idea of muon colliders as tools for particle physics has recently looked promising, there is no example as yet of a successful muon collider.

BUT THERE is an overarching issue of costs that dominates the answer to the question, "How large can accelerators and colliders become, and what energy can they attain?" The relationship of size and cost to energy is determined by a set of relations known as scaling laws. Accelerators and colliders can be broadly classified into linear and circular (or nearly circular) machines. With classical electrostatic accelerators and proton or electron radiofrequency linear accelerators, the scaling laws imply that the costs and other resources required should grow about linearly with energy. Although roughly true, linear scaling laws tend to become invalid as the machines approach various physical limits. The old electrostatic machines became too difficult and expensive to construct when electrical breakdown made it hard to devise accelerating columns able to withstand the necessary high voltages. And radiofrequency linear accelerators indeed obey linear scaling laws as long as there are no limits associated with their required luminosity.

The scaling laws for circular machines are more complex. Ernest

Ernest Courant, M. Stanley Livingston, and Hartland Snyder (left to right), who conceived the idea of strong focussing. (Courtesy Brookhaven National Laboratory)

Lawrence's cyclotrons obeyed an approximately cubic relationship between size or cost and the momentum of the accelerated particle. The magnet's radius grew linearly with the momentum, and the magnet gap also had to increase accordingly to provide enough clearance for the higher radio-frequency voltages required to keep particles and the voltage crest synchronized. All this changed in 1945 with the invention of phase stability by Edwin McMillan and Vladimir Vexler. Their independent work showed that only moderate radio-frequency voltages are required in circular machines because all the particles can be "locked" in synchronism with the accelerating fields.

Then came the 1952 invention of strong focusing, again independently by Nicholas Christophilos and by Ernest Courant, Livingston, and Hartland Snyder (see photograph on the right). Conventional wisdom says that a magnetic lens to focus particles both horizontally and vertically cannot be constructed— in contrast to optical lenses, which can. But the principle of strong focusing showed that, while a magnetic lens indeed focuses in one plane and defocuses in the orthogonal plane, if two such lenses are separated along the beam path, then their net effect is to focus in both planes simultaneously. This breakthrough made it possible to squeeze beams in circular (and also linear!) accelerators to much tighter dimensions, thus reducing magnetic field volumes and therefore costs.

Because the basic linear scaling laws apply to linear machines for both electrons and protons, prominent physicists predicted that all future accelerators would eventually be linear. But the question remained, "Where is the crossover in costs between circular and linear machines?" New inventions, particularly strong focusing, raised the predicted crossover to much higher energy. Moreover, strong focusing also made the scaling law for high energy proton synchrotrons almost linear. The transverse dimensions of the beam aperture do not need to grow very much with energy; thus the cost of large circular proton colliders grows roughly linearly with energy.

While the scaling laws for proton machines are not affected significantly by radiation losses (although such losses are by no means negligible for the largest proton colliders), they become the dominant factor for circular electron machines. The radiation energy loss per turn of a circulating electron varies as the fourth power of the energy divided by the machine radius. It is also inversely proportional to the mass of the circulating particle, which tells you why electrons radiate much more profusely than protons. In an electron storage ring, certain costs are roughly proportional to its radius while others are proportional to the radiation loss, which must be compensated by building large and expensive radio-frequency amplifiers. As the energy grows, it therefore becomes necessary to increase the radius. The total cost of the radio-frequency systems and the ring itself will be roughly minimized if the radius increases as the square of the energy.

Such a consideration therefore indicates that linear electron machines should eventually become less expensive than circular ones. But what is meant by the word "eventually?" The answer depends on the details. As designers of circular electron machines have been highly resourceful in reducing the costs of components, the crossover energy between circular and linear colliders has been increasing with time. But it appears likely that CERN's Large Electron-Positron collider (LEP), with its 28 kilometer circumference, will be the largest circular electron-positron collider ever built.

The only reasonable alternative is to have two linear accelerators, one with an electron beam and the other with a positron beam, aimed at one another—thereby bringing these beams into collision. This is the essential principle of a linear collider; much research and development has been dedicated to making such a machine a reality. SLAC pioneered this technology by cheating somewhat on the linear collider principle. Its linear collider SLC accelerates both electron and positron beams in the same two-mile accelerator; it brings these beams into collision by swinging them through two arcs of magnets and then using other magnets to focus the beams just before collision. In the SLC (and any future linear collider), there is a continuing struggle to attain sufficient luminosity. This problem is more severe for a linear collider than a circular storage ring, in which a single bunch of particles is reused over and over again thousands of times per second. In a linear collider the particles are thrown away in a suitable beam dump after each encounter. Thus it is necessary to generate and focus bunches of exceedingly high density.

An extremely tight focus of the beam is required at the point of collision. There are two fundamental limits to the feasible tightness. The first has to do with the brightness of the sources that generate electrons and positrons, and the second is related to the disruption caused by one bunch on the other as they pass through each other. According to a fundamental physics principle that is of great importance for the design of optical systems, the brightness (by which I mean the intensity that illuminates a given area and is propagated into a given angular aperture) cannot be increased whatever you do with a light beam—or, for that matter, a particle beam. Thus even the fanciest optical or magnetic system cannot concentrate the final beam spot beyond certain fundamental limits set by the brightness of the original source and the ability of the accelerating system to maintain it.

The second limit is more complex. The interaction between one beam and another produces several effects. If beams of opposite charge collide, one pinches the other, usually increasing its density; but if that pinching action becomes too severe, the beam blows up! In addition, the extremely high electric and magnetic fields that arise in the process cause the particles to radiate; the energy thereby lost diversifies the energy of the different particles in the bunch, which makes it less suitable for experiments.

And there is an additional feature that aggravates the problem. As the energy of colliders increases, the cross sections of the interesting reactions decrease as the square of the energy. Therefore the luminosity—and therefore the density of the interacting bunches—must increase sharply with energy. Thus all the problems cited above will become even more severe.

As a result of all these factors, a linear collider is not really linear in all respects; in particular, the brightness of the beam must increase as a high power of its energy. This fact is difficult to express as a simple cost-scaling law. It suffices to say that all these effects eventually lead to a very serious limit on electron-positron linear colliders. Where this limit actually lies remains in dispute. At this time an upper bound of several TeV per beam is a reasonable estimate. We can hope that human ingenuity will come to the rescue again—as it has many times before when older technologies appeared to approach their limits.

THIS DISCUSSION of linear electron-positron colliders is a part of a larger question: "How big can accelerators and colliders, be they for electrons and positrons or for protons, become?" As indicated, the costs of electron-positron linear colliders may be linear for awhile, but then costs increase more sharply because of new physical phenomena. The situation is similar for proton colliders. The cost estimates for the largest proton collider now under construction—CERN's Large Hadron Collider—and for the late lamented SSC are roughly proportional to energy. But this will not remain so if one tries to build machines much larger than the SSC, such as the speculative Eloisatron, which has been discussed by certain European visionaries. At the energy under consideration there, 100 TeV per beam, synchrotron radiation becomes important even for protons and looms as an important cost component. Indeed, physical limits will cause the costs eventually to rise more steeply with energy than linearly for all kinds of machines now under study.

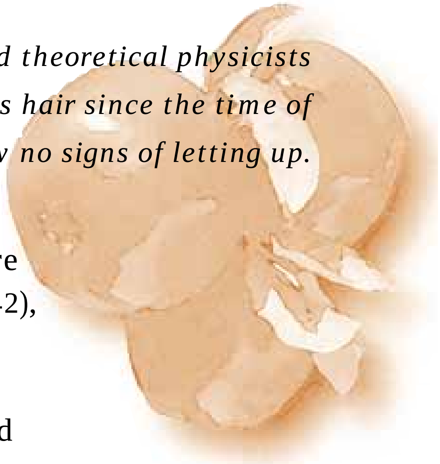
But before that happens the question arises: "To what extent is society willing to support tools for particle physics even if the growth of costs with energy is 'only' linear?" The demise of the SSC has not been a good omen in this regard. Hopefully we can do better in the future.

THE UNIVERSE AT LARGE

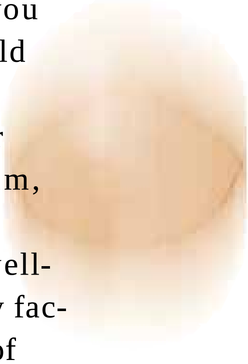
by VIRGINIA TRIMBLE

The Astro-Particle-Cosmo-Connection

Observational astronomers and theoretical physicists have been getting in each other's hair since the time of Newton and show no signs of letting up.



FOR ISAAC NEWTON (1642–1727), though there were laboratory data from the work of Galileo (1564–1642), the British Union of Growers of Poorly-Attached Apples (BUGPAA), and probably others, the real test of universal gravitation was its application to the lunar and planetary orbits that Johannes Kepler (1571–1630) had managed to extract from the observations of his mentor Tycho Brahe (1546–1601). Looking at the various dates, you might reasonably suppose that the planetary orbits would have been somewhat improved by the time *Principia* approached publication (1687), but as the names of other seventeenth-century astronomers will not be on the exam, you are not required to read or remember any of them.



Entering the twentieth century, we find the equally well-known example of Einstein's theory of general relativity facing reality in the form of the advance of the perihelion of Mercury* and the gravitational deflection of light by the sun.** From that day (1919) to this, GR has passed every test astronomy can throw at it, especially the correct description of the changing orbits of binary pulsars (meaning neutron stars in orbits with other neutron stars or massive white

---

*Meaning that Mercury's elliptical orbit rotates once every 3 million years relative to the distant stars.

**Meaning that the apparent positions of stars, and radio sources, have been seen to be shifted on the sky when your line of sight passes close to the limb of the sun.

dwarfs). In the pulsar case, for which Joseph Taylor and Russell Hulse shared the 1993 Nobel Prize in physics, the physical processes include gravitational radiation and other strong-field effects, for which general relativity makes different predictions from those of other theories that would also fit the solar system, weak-field data.

Those pulsar orbits would be getting larger or smaller if the coupling constant, G, were changing with time. Non-zero dG/dt would also affect the lifetimes of stars (whose rate of energy generation scales like G^5), the range of masses possible for old white dwarfs (supported against gravity by degenerate electron pressure) and neutron stars (supported by degenerate neutron pressure), the dynamical evolution of clusters of stars, and distances within the solar system. Curiously, astronomical observations lead to just about the same limits on dG/dt from all of these systems: not more than about 10 percent either way in the 10–20 Gyr age of the universe. Such observations, as well as the Mercurian orbit advance, also tell us that the speed of gravitons is very close to the speed of photons in a vacuum. One always writes the equations with c, but one means c(gravity), not c(light).

OTHER HISTORICAL EXAMPLES AND FALSE ALARMS

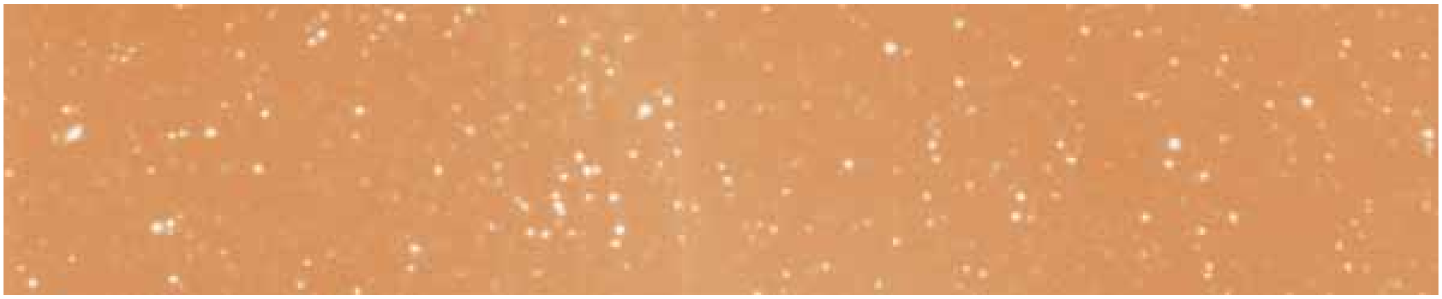
Particle physics can perhaps be said to have begun with the discovery of entities beyond the n, p, and e found in ordinary atoms. The first were the positron, the mu ("Who ordered that?") meson, and the pi (Yukawa particle) meson. All first appeared as upper-atmosphere secondary cosmic rays (ones produced when primary cosmic ray protons hit atmospheric molecules—very hard). A convenient date to remember is 1937, when a shower of papers by people you have heard of in other contexts (Heitler, Oppenheimer, Serber, Homi Bhabha) clarified that these were indeed secondary products but also new particles with well-defined properties.

Astronomical considerations have also made occasional contributions to nuclear physics, most famously in 1953, when Fred Hoyle realized that the carbon-12 nucleus must have a particular excited state, or we would all be made out of pure hydrogen and helium, no further fusion being possible in stars. More recently, the need for lifetimes, energy levels, and cross sections of nuclides likely to form in exploding stars, but most unlikely in the lab, have driven both calculations and experiments.

Russell Hulse, co-discoverer of the binary pulsar 1913 + 16, whose behavior in the decades since has provided the most stringent available tests of general relativity. (It passed; Hulse won a Nobel Prize.) (Courtesy AIP Meggers Gallery of Nobel Laureates)

From time to time, astronomers have concluded that some set of observations simply could not be explained in terms of the (then) standard model of physics and have attempted to invent what they thought was needed. Like many other examples of hubris, this has typically been punished, exile from the community being the most frequent outcome. Some cases are relatively well known, like the Steady State universe, invented to allow stars and galaxies to be older than the apparent cosmic expansion time scale, but requiring the addition of a creation field to general relativity or other theories of gravity. The suggestion that atomic spectral lines can be redshifted by something that is not a Doppler effect, not the expansion of the universe, and not a strong gravitational field, at least when those lines come from quasars, is another well known example.

Less famous, perhaps, are James Jeans' proposal that spiral nebulae represent new matter pouring into the universe, "white hole" explanations of quasars, and the prestellar matter of Viktor Ambartsumian, who believed that clusters of new stars expand out of regions of very



dense, prestellar stuff, perhaps a bit like Gamow's Ylem, but not confined to the early universe, and then in turn expel gaseous nebulae from their surfaces to produce configurations like the stars and gas of Orion. (Conventional stellar evolution tracks do roughly the reverse, beginning with gas and ending with very dense remnants.)

As time goes on, the various possible interactions between astronomy, cosmology, particle physics, and so forth that are discussed in the following sections will move to this one. I am not prepared to guess which will then be seen as "interesting historical examples" and which as "that was an astronomer who thought he was Feynman."*

Academician Viktor Ambartsumian, who died last year, was among the first astronomers to propose a specific mechanism for the formation of expanding clusters of massive, young stars. He later extended the idea (expansion from some kind of very dense, prestellar material, different from known interstellar or laboratory gases) into a possible explanation for quasars. (Courtesy AIP Emilio Segrè Visual Archives)

Leo Goldberg

THINGS THAT DO NOT GO BUMP IN THE NIGHT

You could write a whole book about the exotic particles, properties, and processes that we know do not exist because they would violate some set of astronomical observations. In fact someone has (Georg Raffelt; see the list of "more reading" on page 51). The general idea is that stars must be allowed to form from the interstellar medium, do their nuclear thing for millions or billions of years, and die as planetary nebulae + white dwarfs (from low mass stars) or as supernovae + neutron stars or black holes (from stars of more than about 8 solar masses), at all times adhering to a set of nonlinear differential equations that describe conservation laws, rates of energy generation and transport, and equilibrium between pressure and gravity. The detection of neutrinos from SN 1987A with very much the temperature, time scale, and flux that had been expected from neutron star formation brought this whole field into considerable prominence. The constraints are sometimes quite tight simply because, on the whole, stars manage pretty well with just the standard-model physics that we are all so tired of.

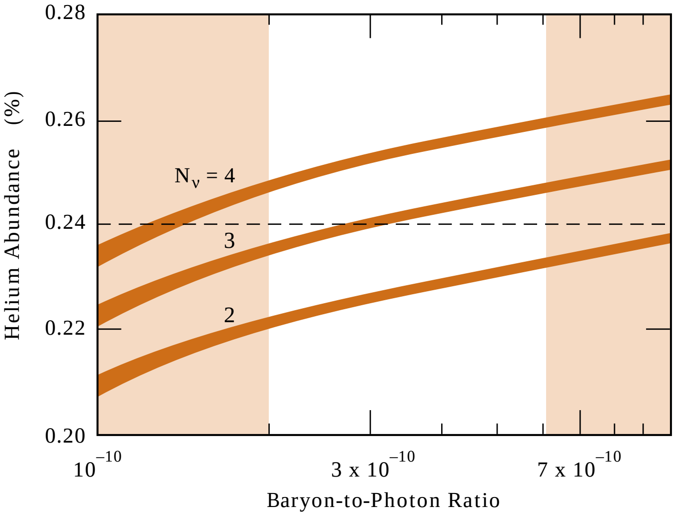
In addition, any new entities you might want to postulate must not be so numerous and massive as to make the average density of the universe big enough to slow the expansion measurably today (since we see no such slowing). Nor are they (or you) allowed to spoil the set of nuclear reactions at high density and temperature that produce deuterium, helium (3 and 4), and a bit of lithium-7 in the early universe ("big bang nucleosynthesis"). I mention here only a small, representative set of examples and urge you to peruse Raffelt's book for many more and for the corroborative details.

1. There must not be too many magnetic monopoles floating around, or they will short out the large-scale magnetic fields of Jupiter, pulsars, and the interstellar gas. This Parker (for Eugene of Chicago) limit means that such monopoles must have rest masses of at least 10^{16} GeV if they are to be dynamically important in the universe.

2. The magnetic dipole moment of the electron neutrino cannot be more than about 3×10^{-12} of the Bohr magneton, or the neutrinos from SN 1987A would never have

---

**The original image here was a* New Yorker *cartoon bearing the caption: "That's God. He thinks he's a doctor."*



Plot of some of the consequences of nucleosynthesis during the hot dense (big bang) phase. Observed abundances of lithium-7 and deuterium in gas and stars that have experienced very little nuclear processing require that the real universal density of baryonic material (the baryon-to-photon ratio) fall somewhere in the white stripe—corresponding to a baryon density less than 10 percent of the closure density. Then the fact that the abundance of helium is, at very most, a little more than 24 percent says that there can be at most three neutrino flavors in the early universe. (Courtesy C. Copi and D. Schramm, University of Chicago)

made it out of the parent star and through space to us. This is probably rather smaller than the best laboratory limit, and the 1987A data also set limits to neutrino masses, coupling to right-handed and Majorana neutrinos, and such that are comparable to or better than the laboratory numbers.

3. Quite a few of the things you might think of doing with neutrinos would mess up the early universe, including, in particular, adding to the "known" three flavors. A fourth or fifth neutrino family would speed up the early expansion so much that too many neutrons would survive to form more helium than we see. There is, of course, also a limit of roughly three neutrino flavors from the laboratory width of Z^0 decay, but, because we do not know lifetimes or masses a priori, the two considerations rule out somewhat different volumes of parameter space.

4. Any new bosons or weakly interacting massive particles you might want to dream up must not couple to ordinary or degenerate matter tightly enough to transport much energy in either normal stars or white dwarfs and neutron stars. If they do, you will cool off your WDs and NSs too fast (so we wouldn't see the ones we see) and change the internal density and temperature distribution of nuclear-burning stars away from the ones needed to reproduce known correlations of stellar masses, luminosities, radii, and evolutionary phase.

A cross section of 10^{-36}cm^2 at stellar temperatures borders on being "too big" for a number of these contexts. Another false alarm was the attempt to reduce neutrino emission from the sun by cooling its interior with WIMPs whose cross sections fell in the borderline range. At least two problems resulted. The interior distribution of density no longer matched the one derived from analysis of solar pulsation frequencies, and later stages of evolution, like the horizontal branch phase, became so short-lived that you couldn't account for the large numbers of stars seen in them.

There are also a few cases where something new under the sun might still improve agreement between models and observations. One of these is the possible presence of pion condensate or strange quark matter in the interiors of neutron stars (which we should then call pion stars, quark stars, or some such). Either one will hasten cooling after nuclear reactions stop. This could be useful if existing upper limits on thermal emission from the surfaces of neutron stars should ever get pushed lower than the predictions from conventional cooling curves. In addition, each permits a given mass to be somewhat more compact without collapsing. Thus the star can rotate a bit faster without slinging mud in theorists' faces. At the moment (2:37 p.m. Wednesday, September 25, 1996) the two shortest periods of rotation measured for neutron stars are both quite close to 1.55 msec and are comfortably accommodated by most ordinary equations of state for nuclear matter. The false alarm of a 0.5 msec pulsar reported at the site of SN 1987A several

years ago triggered a considerable flurry of preprints on quark (etc.) stars, some of which made it into print before the report was retracted—and a few afterwards!

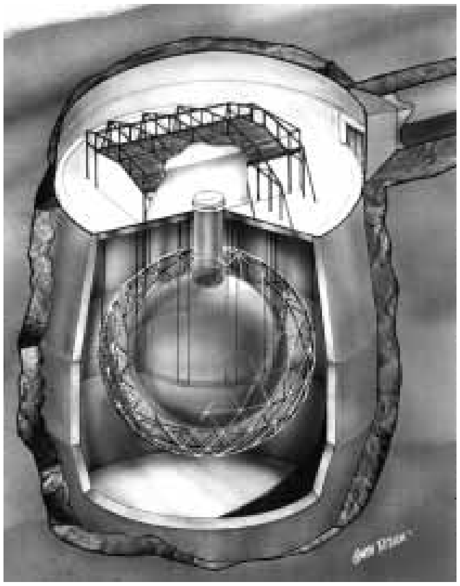
Neutron stars remain, of course, the most extreme environment under which we can test pictures of how superfluids and superconductors behave. They also remain awkwardly refractory to experiment.

THERE'S GOT TO BE A PONY IN THERE SOMEWHERE*

The two topics on which nearly everybody agrees that astronomers and particle physicists must cooperate if answers are ever to be found are "the solar neutrino problem" and the complex of questions concerning the existence and nature of dark matter, the origin of large-scale structure in the universe (formation and distribution of galaxies and clusters of galaxies), and whatever happened before big bang nucleosynthesis, including inflation, baryogenesis, phase transitions, and miracles. Neither is at all new to regular readers of these pages.

John Bahcall summarized the solar neutrino situation here (see the Fall/Winter 1994 *Beam Line*, Vol. 24, No. 3, page 10). I will summarize still further. First, Raymond Davis Jr.'s chlorine-37 experiment has been seeing a bit less than a third of the predicted flux of high energy neutrinos since before 1970, and the first generation of possible excuses already included many of the astronomical and weak-interaction fiddles that are still with us (for examples see the Trimble and Reines review mentioned under "more reading"). Second, three additional experiments have not clarified things as much as one might have hoped. At the very highest energies that come only from boron-8 decay, the Kamiokande electron-scattering detector has reported about half the number of expected events from the direction of the sun (none of the other devices provides any directional information). And the SAGE and GALLEX gallium detectors also see about half the expected flux, mostly in the form of lower energy neutrinos from the proton-proton reaction ($p + p \rightarrow d + e^+ + \nu_e$).

Third, it is rather difficult to make this combination come out from any fiddle you can think of, mostly because it is the middle energy range that seems to be most deficient. New weak interaction physics, along the lines of neutrino oscillations catalyzed by the presence of nuclei (MSW effect), seems to work better than non-standard models of the solar interior. Fourth, even MSW-type oscillations are squeezed into a very narrow corner of the space of neutrino masses and coupling constants when you also insist on accounting for the anomalous ratio of neutrino flavors among cosmic-ray secondaries made in the atmosphere. Fifth, new detectors under construction or planned (SNO, SuperKamiokande, Borexino) could sort things out (but need not), and I suspect that the last word has not been said on this topic, not even my last word.



Artist's conception of the Sudbury Neutrino Observatory (SNO) detector. When fully operational, it will detect all three flavors of neutrinos and give some indication of the direction from which they come. Although sensitive only to the very highest energy (boron-8) solar neutrinos, it should be able to decide if some of the missing electron neutrinos have rotated into mu- or tau-neutrinos.

(Courtesy Lawrence Berkeley National Laboratory)

Readers who remember the joke of which this is the punch line are invited to share it with those who don't, preferably keeping in mind that roughly half the preprint pile comes from my side of the interdisciplinary fence and half from yours—unless we are on the same side.

Finally, we come to the constellation of issues associated with dark matter and the very early universe. The observational situation is quickly summarized: 90 percent or more of the stuff in the universe that contributes to gravitational potentials does not emit (or absorb) its fair share of electromagnetic radiation. Dark matter unquestionably exists and outweighs the luminous matter in stars, galaxies, and the gas between them. But we haven't a clue what it is.

Colleagues often object to this second statement. What they mean, however, is not that we have any very definite information about what the dark matter is, but only that we know quite a lot of things it is not. This is progress only if the number of ideas generated by theorists is finite (not by any means a safe bet). For starters, the requirement of not messing up big bang nucleosynthesis almost certainly means that the dark matter cannot all be ordinary stuff made of protons, neutrons, and electrons. Thus we are forced to hypothesize other stuff that is capable of, at most, gravitational and weak interactions, and not of electromagnetic or nuclear ones (again a few colleagues would disagree at some level).

Dark matter, structure formation, inflation, phase transitions, etc. get mixed up together in several ways. First, most obviously, galaxies and clusters live in potential wells made mostly of dark matter, and the nature of the stuff is bound to make a big difference to how galaxies form (and whether we can model them at all successfully, to which the present answer is no, not entirely). Second, galaxy formation might be aided (or impeded) by various topological singularities (cosmic strings, textures, . . .) left from the phase transitions associated with the four forces gradually separating themselves. The supersymmetry arguments that go with the forces having once been the same more or less automatically imply the existence of several kinds of nonbaryonic particles associated with assorted unfamiliar but conserved quantum numbers.

Third, the "inflaton field" responsible for early, exponential expansion of the universe (inflation) could possibly leave behind a small ghost of itself to act as a cosmological constant (Einstein's unloved Λ). Fourth, inflation, at least some kinds, is supposed to leave behind both the exact critical density required to stop universal expansion in infinite time and a spectrum of perturbations of that density with a definite form, well shaped to grow into galaxies and clusters. No obvious astronomical observation would seem capable of proving that inflation happened, but one could imagine definitive dynamical evidence for a total density less than the critical one or for a spectrum of not-yet-evolved density perturbations different from the inflationary prediction. But there are already variants of inflation in the literature that can live with one or both anomalies.

In some ways, this mess looks slightly simpler from the astronomical side. As far as we can tell, for the purposes of galaxy formation and creation of large-scale structure, everything nonbaryonic can be divided among four categories, and it doesn't much matter which example nature has chosen to favor. The four categories are non-zero cosmological constant, seeds (like the topological singularities), hot dark matter (consisting of particles light enough that they are relativistic at $T \approx 3000$K when baryonic matter and light stop talking to each other; ordinary neutrinos of 5–25 eV are the most obvious candidate), and cold dark matter (consisting of particles massive enough to be non-relativistic at the same temperature, like the lowest-mass supersymmetric particle and its cousins; or axions which are low mass but form at rest; and no, I don't know why).

You can, if you wish, have two of these or even three. I am not aware of any scenarios that involve all four simultaneously, but this may well come. The variety is welcomed because no current simulation of galaxy (etc.) formation simultaneously does a very good job of accounting for structures on relatively small linear scales (a megaparsec or less, promoted by CDM), the largest scales (up to 100 Mpc, promoted by HDM), the largest deviations from smooth cosmic expansion that we see, and the observed sizes of those deviations (for example, the dispersion of pair-wise velocity differences between nearby galaxies) as a function of scale length. Choosing a spectrum of initial density fluctuations different from the standard inflationary one allows yet another degree

of freedom. It is not, I think, clear whether what is needed is just further exploration within the territory described above or whether there may still be some important piece of physics missing from the simulations.

There is, however, one thing you can be sure of. I am not going to be the person to holler that the astronomical observations require new physics (or new imperial clothes, or whatever) or to suggest the form that physics should take.

MORE READING

For the multitude of limits on particle properties that arise from considerations of stellar structure, see G. G. Raffelt, **Stars as Laboratories for Fundamental Physics,** 1996, University of Chicago Press.

Strange Quark matter is discussed in G. Vassiliadis et al. (eds) **Proc. Int. Symp. Strangeness and Quark Matter**, World Scientific Press, Singapore and in *Nuclear Physics B* (Proc. Supplement) 24B on Strange Quark Matter in Physics and Astrophysics, 1992.

Atmospheric neutrinos are featured in T. K. Gaiser et al. (1995) *Phys. Reports* 258, 173 and in M. Fukugita and A. Suzuki (Eds.) 1994, **Physics and Astrophysics of Neutrinos** (Springer-Verlag).

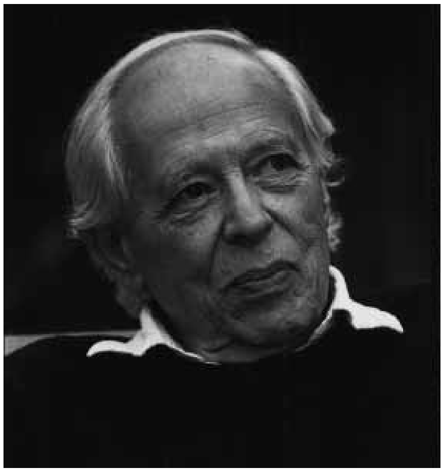
Various snapshots of the solar neutrino problem appear in V. Trimble and F. Reines, 1973, *Rev. Mod. Phys.* 45, 1; J. N. Bahcall, **Neutrino Astrophysics** (1989), Cambridge University Press; and Y. Susuki and K. Nakamura (Eds.) 1993, **Frontiers of Neutrino Astrophysics** (Universal Academy Press, Tokyo).

For the various kinds of WIMPs, inos, and other dark matter candidates implied by supersymmetry, see G. Jungman, M. Kamionkowski, and K. Griest 1995, *Phys. Reports.*

And, finally, inflation and other highlights of the early universe appear in

A. Linde 1991, **Particle Physics and Inflationary Cosmology**, Harvard Univ. Press, E. W. Kolb and M. S. Turner 1990, **The Early Universe**, Addison-Wesley, and G. Boerner, **The Early Universe, Fact and Fiction**, 2nd ed. 1992, Springer-Verlag.

CONTRIBUTORS


J. Selsing

ABRAHAM (BRAM) PAIS is the Detlev W. Bronk Professor Emeritus of The Rockefeller University. Born in Amsterdam in 1918, he earned his Ph.D. degree from the University of Utrecht in 1941 and emigrated to the United States after World War II. In the 1950s he introduced the concept of associated production, which governs the behavior of strange particles, and developed the idea of particle mixing. He resides in New York City and Copenhagen, Denmark.

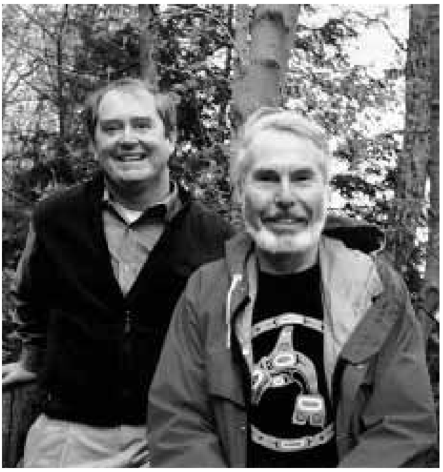
Among his many publications is *Subtle is the Lord: The Science and Life of Albert Einstein*, which in 1983 won the American Book Award for Science and the American Institute of Physics Science Writing Award. His other books include *Inward Bound*, *Niels Bohr's Times* and *Einstein Lived Here*. For his extensive contributions to the public understanding of science, he received the 1993 Gemant Award of the American Institute of Physics and the 1995 Lewis Thomas Prize.


L. Weinberg

STEVEN WEINBERG is a member of the Physics and Astronomy Departments at the University of Texas, and the founding director of its Theory Group. His work in physics has been honored with the Nobel Prize, the National Medal of Science, the Heinemann Prize in Mathematical Physics, 13 honorary doctoral degrees, and election to the National Academy of Sciences, the Royal Society, the American Philosophical Society, and the American Academy of Arts and Sciences.

He is the author of over 200 scientific articles, one of which is the most cited article of the past 50 years in particle physics, and seven books of which the most recent is *The Quantum Theory of Fields*.

Educated at Cornell, Copenhagen, and Princeton, he taught at Columbia, Berkeley, MIT, and Harvard before coming to Texas in 1982.


L. Quigg

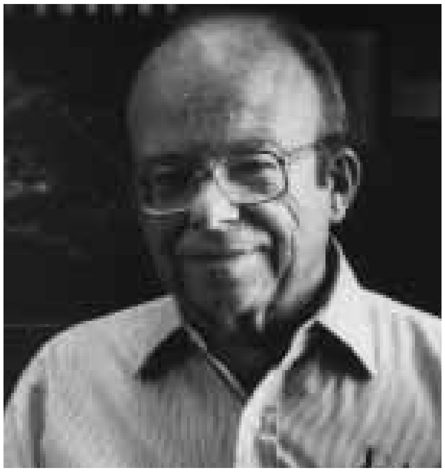
CHRIS QUIGG, left, is a member of the Theoretical Physics Department at Fermilab and Visiting Professor at Princeton University. His Ph.D. research at Berkeley was more distant from quarks and gauge fields than today's students can possibly imagine. It nevertheless began his lifelong engagement with experiment and his close association with J. D. Jackson (right), who is teaching him still.

A recurring theme in Quigg's research is the problem of electroweak symmetry breaking and the exploration of the 1 TeV scale. His current interests include mesons with beauty and charm, the top quark, and neutrino interactions at ultrahigh energies. He is also at work on a second edition of *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*.

E. Heitowit

MICHAEL RIORDAN has spent much of his last twenty years writing and editing general books on science, technology, and their impact on our lives. They include *The Solar Home Book* (1977), *The Day After Midnight* (1981), *The Hunting of the Quark* (Simon & Schuster, 1987), *The Shadows of Creation* (W. H. Freeman, 1991) and the forthcoming *Crystal Fire: The Birth of the Information Age* (W. W. Norton, 1997)—from which his article in this issue has been adapted.

He currently divides his working time between SLAC, where he serves as Assistant to the Director and Contributing Editor of the *Beam Line*, and the Santa Cruz Institute for Particle Physics, where he is researching a scholarly history of the Superconducting Super Collider. For recreation he can often be found paddling his kayak on the waters of Monterey Bay or hiking near his home in the Santa Cruz Mountains.



WOLFGANG (PIEF) PANOFSKY is Professor and Director Emeritus, Stanford Linear Accelerator Center. A member of the National Academy of Sciences, he has served as Chairman of its Committee on International Security and Arms Control and currently heads its Weapons Plutonium Management and Disposition Study Committee. He served on the President's Science Advisory Committee in the Eisenhower and Kennedy Administrations.

Born in Berlin in 1919, Pief earned his Ph.D. in 1942 from Caltech. After the war, he joined the University of California at Berkeley, then came to Stanford as Professor of Physics in 1951. He served as Director of its High Energy Physics Laboratory from 1953 to 1961 and was the founding Director of SLAC, continuing in that position until his "retirement" in 1984. Among his many awards are the National Medal of Science and the Fermi and Lawrence Awards of the Department of Energy.

"The Universe at Large" appears this issue largely thanks to the author's husband, Dr. Joseph Weber (on her left), who took her out to dinner every evening for a week to free up some time for writing it. They continue to oscillate between the University of Maryland and the University of California at 31.7 nHz, partly because of the interestingly different sets of restaurants on the two coasts. The chap in front is Yakov B. Zeldovich, with whom they had a very memorable dinner in Jena in 1982. He was among the pioneers in the USSR in looking at the interface between particle physics and cosmology.