The impact of special relativity on theoretical physics

Rudely uniting absolute space and universal time into a single, changeable 'space–time,' the theory has become a fact of life, part and parcel of our view of nature; this essay illustrates its impact by means of some early and late examples of its use.

J. David Jackson

As one of my colleagues put it, "Asking about the impact of special relativity on theoretical physics is like asking about the impact of Shakespeare on the English language." An impossibly large, even senseless, task. Special relativity is a fact of life, part and parcel of the way nature is. If its impact on everyday life is slight, its impact on physicists' thinking is profound. Space and universal, inexorable time were rudely united into space–time by Albert Einstein's discovery of special relativity. Explorations of the properties of space–time, of the covariance of physical laws, and of the physical invariants of nature, together with quantum mechanics, led to the formulation of relativistic quantum field theories. The external symmetries of space–time were augmented by abstract spaces for internal symmetries corresponding to invariant (conserved) quantities such as isospin, strangeness and charm. The whole worldview of modern theoretical physics can be traced back to the fundamental postulate or idea that physical phenomena do not change just because you happen to be moving by, instead of standing still, when observing them. Rotations and other transformations in internal spaces have replaced "moving by," but the idea is the same.

The idea that the laws of nature, at least on a scale that does not encompass the whole universe, are independent of translational motion of the system under study is very old—certainly it was part of Galileo's knowledge. In special relativity, the idea is elevated to be the first postulate:

▶ There exists a triply infinite set of equivalent Euclidean frames of reference, moving with constant velocities in rectilinear paths relative to one another, in which all physical phenomena occur in an identical manner.

The second postulate is the joker, simplicity itself:

▶ There exists in nature a limiting, invariant speed.

The existence of a limiting, invariant speed is a crucial part of special relativity; it happens to be the speed of light, but it is shared by all massless particles and approached very closely by cosmic rays and by ordinary particles in accelerators.

For many physicists at the turn of the century special relativity was intimately intertwined with electromagnetism and light—not surprisingly, since swiftly moving light and slowly moving matter seem so distinct, and since the historical roots of relativity lay in attempts to reconcile electromagnetic experimental results with a mechanistic ether. Einstein did not help by using as his second postulate "light is aways propagated in empty space with a definite velocity c, independent of the state of motion of the emitting body," or by his use of light signals to synchronize his clocks, or by titling his 1905 paper "On the electrodynamics of moving bodies." That the laws of relativity are very general was recognized by Wolfgang Pauli[1] in 1919 and by others even earlier.

From the two postulates, plus some reasonable assumptions about the isotropy of space–time and such, flow the consequences of special relativity, with its demotion of time to just another coordinate and its blurring of the distinction among before, now and after. Although there are many seemingly more bizarre consequences of special relativity, it was the destruction of the preferred position of time that made relativity difficult for so many in the early days and even now. All of this is old, familiar stuff. The broad sweep from 1905 to the present day is also a well-trodden path.[2]

Einstein was apparently familiar with the Michelson–Morley experiment, published 18 years earlier, when he wrote his paper on special relativity, even though he does not refer to it explicitly. (See the article by John Stachel on page 45.) Whether Einstein acknowledged it or not, the Michelson–Morley experiment was a very important component of the general milieu of theoretical physics in 1905. It certainly played a central role in the efforts of Hendrik A. Lorentz and others to

J. David Jackson is professor of physics at the University of California, Berkeley. He is a particle theorist and the author of a well-known text on classical electromagnetism. He has recently been working on the SSC.

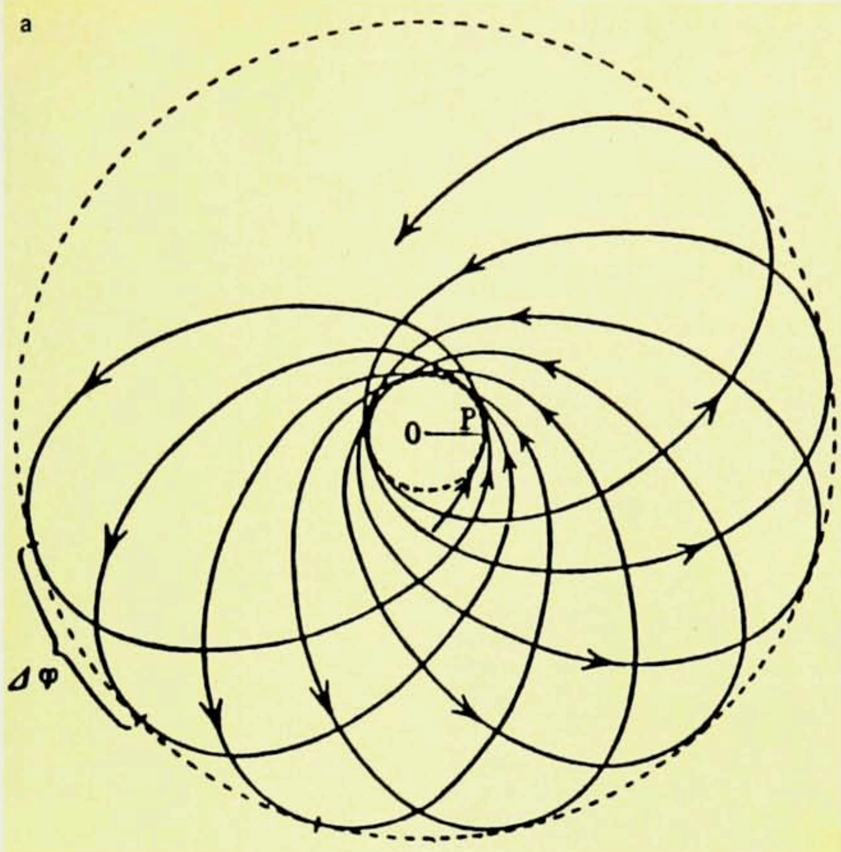
0031-9228 / 87 / 0500 34– 09 / $01.00

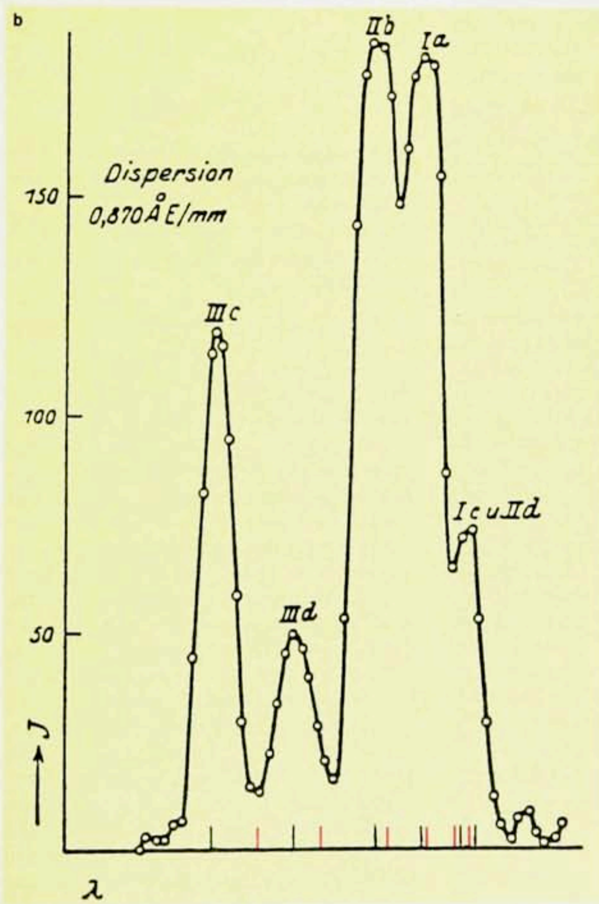
develop a consistent electrodynamic theory of matter within the framework of the Maxwell equations and an ether. But with his 1905 paper Einstein changed the rules of the game. In one stroke, he made the ether meaningless, generalized classical mechanics to relativistic speeds and comfortably ensconced an etherless Maxwell's theory within his framework. Small wonder that his peers balked for a few moments at swallowing such a large pill!

This informal essay is in the nature of a lecture written down, with the impact of special relativity on theoretical physics illustrated—rather superficially, I fear—by examples of its use. Some are well known, others perhaps less familiar. The examples are taken largely from the first 30 years after Einstein's discovery. In recent times, the theory has so permeated theoretical particle physics that it is pointless to describe its effect. Instead, I give a few illustrations of the efficacy of viewing a physical phenomenon in different reference frames, thus emphasizing that there is no preferred inertial frame and returning the discussion to the theme of this issue, the Michelson–Morley experiment.

Electrons and atoms

One test of special relativity is to verify the expressions for energy and momentum as functions of speed for a particle, such as the electron—recently discovered when Einstein devised his theory. By 1905, this test was already





Hydrogen atom. Above is Arnold Sommerfeld's diagram for the orbit of an electron in a hydrogenlike atom. Relativistic corrections perturb the perfect Keplerian orbits of the Bohr model, leading to splittings between energy levels seen as fine structure in the spectra of hydrogenlike ions. (From reference 4.) The graph at left shows the fine structure of the He^+ spectrum obtained by Fritz Paschen (from *Ann. Phys., Leipzig*, **50**, 901, 1916). The data are photometer readings of a photographed spectrum in the vicinity of the 4686-Å line ($n = 4 \rightarrow n = 3$); the scan covers 0.6 Å. The tick marks on the abscissa are the predictions of Sommerfeld's model, in perfect agreement with the data; we have added (in color) marks indicating the predictions of Max Abraham's "absolute theory."

under way, thanks to the electrodynamic models of the electron of Max Abraham, Lorentz and others. Abraham's model was a rigid sphere of charge; for a moving electron it gave the relations

$$E = E_0\left[\frac{1}{\beta}\ln\left(\frac{1+\beta}{1-\beta}\right) - 1\right]$$
$$cp = E_0\left[\frac{1+\beta^2}{2\beta^2}\ln\left(\frac{1+\beta}{1-\beta}\right) - \frac{1}{\beta}\right]$$

where β is v/c and E_0 is the energy of the sphere at rest, $e^2/2a$. Experiments done in 1901–02 by Wilhelm Kaufmann agreed well with Abraham's expressions. By 1904, Lorentz had proposed a model in which the electron's charge is subject to the FitzGerald–Lorentz contraction. His formulas were

$$E = E_0\gamma(1 + \beta^2/3)$$
$$cp = E_0'\gamma\beta$$

where γ is the familiar Lorentz factor of special relativity, $(1-\beta^2)^{-1/2}$. The factors E_0 and E_0' are different. Lorentz concluded that Kaufmann's data were as consistent with his formulas as with Abraham's.

Meanwhile, Jules Henri Poincaré, building on Lorentz's work but removing, at least formally, certain inconsistencies (electrostatic instability, for example), arrived in 1905, and more fully in 1906, at the expressions

$$E = \gamma mc^2$$
$$p = \gamma mv$$

Einstein obtained the same relations, at the same time, on purely kinematic grounds. These are the well-tested and familiar expressions of today.

But model builders die hard, particularly if there are experimenters aiding and abetting. Abraham was in battle against Lorentz and Poincaré. In early 1906 Kaufmann announced a new set of measurements and declared that the Poincaré–Einstein formulas were definitely ruled out by experiment. A. H. Bucherer entered the fray with his own model and his own observations. By 1908, Bucherer, on the basis of new data, had abandoned his model and accepted the standard expressions, but the dust did not fully settle until almost 1920—with all the theorists (except Einstein, who floated aloof above it all) agonizing as the experimental sands shifted around them.[3] It seems that then, as now, the scars of such battles where reputations are on the line color the participants' attitudes about broader issues for all time.

A brilliant example of the full exploitation of special relativity in the early years is Arnold Sommerfeld's calculation of the fine structure of hydrogenlike atoms in the old quantum theory. In late 1915, W. Wilson and Sommerfeld independently generalized Niels Bohr's quantization condition to systems with more than one degree of freedom. The extension is expressed in terms of generalized momenta p_i and coordinates q_i:

$$\oint p_i \,\mathrm{d}q_i = 2\pi n_i \hbar$$

where n_i is an integer and the integration is over a full cycle of the motion of the ith degree of freedom.

Sommerfeld applied the extended formalism to the nonrelativistic hydrogen atom (elliptical orbits), introducing radial (n_r) and angular (k) quantum numbers for the two degrees of freedom in the plane and showing that the Bohr energy formula still applied, provided Bohr's quantum number n is replaced by $n_r + k$. The presence of only the sum of the two quantum numbers, not n_r or k separately, results in the degeneracy (in a nonrelativistic description) of the energy levels with respect to orbital angular momentum, a result that is preserved in quantum mechanics proper. Sommerfeld then moved on to a relativistic description of the atom. He noted that the familiar kinetic energy $p^2/2m$ is only an approximation to the special relativistic expression

$$T = c\sqrt{p^2 + m^2c^2} - mc^2$$

where m is the rest mass of the particle. The Hamiltonian for a hydrogenlike atom then is

$$H = c\sqrt{p^2 + m^2c^2} - mc^2 - \frac{Ze^2}{r}$$

Sommerfeld saw that angular momentum, suitably defined (in terms of polar coordinates) as $\gamma mr^2\dot{\theta}$, is conserved and so is still quantized as an integral (the quantum number k) multiple of $\hbar$. Separation of p^2 into angular and radial parts led to

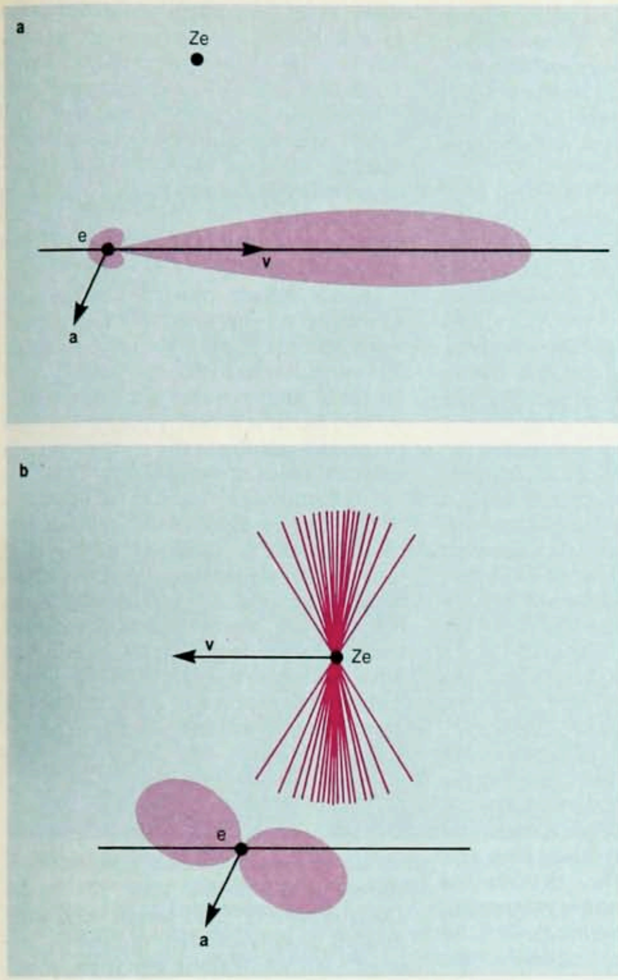
$$H = c\left(p_r^{\,2} + \frac{k^2\hbar^2}{r^2} + m^2c^2\right)^{1/2} - mc^2 - \frac{Ze^2}{r}$$

For constant energy E (the quantity sought for), the expression

$$H = E$$

is a function of r alone and can be solved for p_r in terms of r and constants. Sommerfeld then used the quantization rule to find the energy eigenvalues. In this way, he found the

Llewellyn Hilleth Thomas showed that a precessional effect in special relativity produces an additional energy of interaction of the spin of an orbiting particle. Born in London in 1903 and educated at Cambridge, Thomas taught at Ohio State University and Columbia University. During World War II he worked at the Aberdeen Proving Ground in Maryland. He is now retired from North Carolina State University. (Photo courtesy of AIP Niels Bohr library.)



Bremsstrahlung as seen in two reference frames. **a:** In the laboratory frame, the heavy nucleus (charge *Ze*) is at rest while the swift, light particle (charge *e*) shoots by, emitting a spray of radiation in the forward direction. **b:** In the rest frame of the light particle, the electromagnetic field of the nucleus appears as a pulse of radiation that is scattered as dipole radiation by the light particle.

rather complicated expression[4]

$$E = mc^2\left[1+\frac{\alpha^2Z^2}{(n_r+\sqrt{k^2-\alpha^2Z^2})^2}\right]^{-1/2} - mc^2$$

where α is Sommerfeld's fine structure constant, $e^2/\hbar c$, approximately $1/137$.

The relativistic corrections for the hydrogen atom ($Z=1$) are tiny. Because α is so small, the square root in the denominator above is nearly equal to k. By expanding the overall square root in powers of α^2, one recovers the Bohr formula

$$E = -\frac{\alpha^2mc^2}{2(n_r+k)^2}$$

But the amazing thing is that Sommerfeld's result, with integer k and n_r, agreed to fantastic precision with Friedrich Paschen's careful observations of hydrogenlike atoms such as He^+ (see the figure on page 35). Einstein and many others were impressed. With reduced-mass corrections included, the equation withstood the assaults of greater and greater precision until 1947, when Willis E. Lamb and R. C. Retherford observed the minute splitting between the $2s_{1/2}$ and $2p_{1/2}$ states in hydrogen caused by radiative corrections.

Some of you knowledgeable about quantum mechanics and spin will say, "But fine structure is caused by spin–orbit coupling, not just the relativistic increase in mass! We treated those separately in our quantum mechanics classes." Others will say, "But Sommerfeld got the degeneracies wrong. His k quantum number is really $j+1/2$ and states with the same j but different orbital angular momenta are still degenerate." My answer is that spin–orbit coupling is a relativistic effect and the fact that Sommerfeld obtained precisely the energy formula found 12 years later with the Dirac equation, even if the most detailed correspondences are not quite right, shows that there is a deep connection between special relativity and the concept of spin. I prefer to think that Sommerfeld was somehow prescient, not just lucky. His relativistic treatment within the old quantum theory captured the essence of the spinning electron without knowledge of that degree of freedom!

An interesting sidelight of Sommerfeld's work on fine structure is how he used the experimental data and his quantization methods in support of Einstein's theory.[4] He noted that the relativistic kinetic energy of the electron in the various electrodynamic models and special relativity could be written to first order in the common form

$$T = (p^2/2m)(1+3a\beta^2/4+\cdots)$$

where a is 1 for special relativity and $4/5$ for Abraham's model, for example. He then calculated the first-order correction to the energy levels and arrived at the same expression as given above, except that on the left E is multiplied by a, as are both factors of α^2Z^2 on the right. Evidently, the factors of a are so positioned that in the limit of small α^2Z^2 we recover the Bohr formula, but the fine-structure spacings are proportional to a. Sommerfeld then compared the formula with the experimental spacings and concluded that the data conclusively rule out a value of a of $4/5$ but are in complete agreement with $a=1$. (See the figure on page 35.) Sommerfeld felt, and I do too, that disproof of the "absolute theory" (as he called it) and support of special relativity by precise spectroscopic measurements was a nice touch.

Waves and particles

My next examples of the impact of special relativity on theoretical physics—really on physics—are two sides of a coin: the particulate nature of light (Compton effect, 1922) and the wave nature of matter (the de Broglie hypothesis, 1923). Both involved special relativity intimately, and had profound impact on physical thought. They can, I suppose, be described as second-order effects of special relativity. They are also so well known as to require only brief mention.

Recall that Arthur Holly Compton reported seeing in the scattering of x rays a secondary line at longer wavelengths than the incident radiation, a line whose wavelength was a function of angle. The explanation lies in the conservation laws of relativistic energy and momentum for a collision between a massless photon (x ray) of energy E and momentum E/c and a stationary electron of mass m_e. The scattering act transfers some energy and momentum to the recoiling electron. The energy E' of the scattered photon is hence less than the energy of the incident photon, so its wavelength is longer. Direct application of the formulas of special relativity leads to the Compton relation

$$\Delta\lambda = \frac{2\pi\hbar}{m_e c}(1-\cos\theta)$$

It is striking that the Compton effect, which firmly established the *particle* nature of light quanta (objects carrying both energy and momentum and behaving as particles in relativistic kinematics), has a formula displaying the *wave* nature of the electron through $\hbar/m_e c$. We even call it the Compton *wavelength*!

In his PhD thesis, Louis de Broglie argued that if a light wave has a particulate nature that can be described by an energy–momentum (E/c,**p**), related to the frequency–wavevector (ω/c,**k**) by

$$E = \hbar\omega$$
$$\mathbf{p} = \hbar\mathbf{k}$$

then material particles should have associated with them waves and wavelike behavior with the same connections. Thus he predicted that particles of momentum p should show diffraction phenomena of wavelength $2\pi\hbar/p$, a prediction triumphantly confirmed by experiment in 1926. De Broglie made his whole discussion explicitly relativistic to describe light and matter waves in the same framework. He showed that requiring the phase advance of a wave following a semiclassical closed path to be an integral multiple of 2π yielded the relativistic Wilson–Sommerfeld quantization rules. He thus had relativistic wave mechanics almost within his grasp. The relevant point for us is his exploitation of special relativity to press the particle–wave duality for matter as much as for light.

As an aside, I note a bit of a puzzle: Einstein, the inventor of special relativity and of the light quantum for the photoelectric effect, took about 10 years to admit in print that a light quantum has directed momentum as well as energy associated with it.[5] It seems amazing that someone who was aware of the four-vector character of both (ω/c,**k**) and (E/c,**p**) and who had invented a bundle of light energy with $E = \hbar\omega$ did not get around to accepting $\mathbf{p} = \hbar\mathbf{k}$ for 11 years and did not explain the Compton effect in advance. Perhaps it is a measure of how ingrained the continuous wave theory of James Clerk Maxwell was, and also of how small a photon's momentum E/c is for a given energy as compared with that of a nonrelativistic particle, whose momentum times c is the geometric mean of the kinetic energy and the rest energy, $(2mc^2T)^{1/2}$. Only with energetic photons (x rays and above) does one easily see the transfer of momentum.

Spin and Thomas precession

My next example of how special relativity spread in influence concerns Llewllyn Hilleth Thomas, of the Thomas precession and the sector-focusing cyclotron. The story of the alkali doublets, the anomalous Zeeman effect and the introduction of the spin of the electron by George Uhlenbeck and Samuel Goudsmit in 1925 is familiar to all.[6] But the assumed g-factor of 2 for the electron, necessary for the Zeeman effect, was an embarrassment in the fine structure, where a straightforward calculation of the spin–orbit interaction gave splittings a factor of 2 too large. The location of the error and its elimination to yield the correct fine structure are due to Thomas. It is unfortunate that most physicists are aware only of the short letter[7] Thomas published in *Nature* in early 1926, if of that. It baffled even the greats like Pauli for a time. There is also a longer paper[8] by Thomas that is a marvel to behold. His contribution concerning the fine structure is sophisticated and profound; but not only does the detailed paper contain the Thomas precession, it has the complete development of the relativistically covariant equations for the motion of a spin with any g-factor in an arbitrary external electromagnetic field.

The effect discovered by Thomas is purely kinematic and results from the lack of commutativity of successive Lorentz transformations as a particle moves in a curved path. He showed that the coordinate axes in the particle's rest frame precess with respect to the laboratory axes with an angular velocity

$$\boldsymbol{\omega}_T = \frac{\gamma^2}{\gamma+1}\frac{\mathbf{a}\times\mathbf{v}}{c^2}$$

where γ is the usual Lorentz factor and **a** is the acceleration. For a particle of spin **s**, this Thomas precession adds a term to the potential energy equal to $\mathbf{s}\cdot\boldsymbol{\omega}_T$. For a nonrelativistic electron moving in a central electrostatic potential energy $V(r)$, the simple spin–orbit energy (caused by the electric field appearing partially as a magnetic field in the electron's rest frame) and the Thomas precession energy neatly combine to give

$$U = \frac{g-1}{2mc^2}\mathbf{s}\cdot\mathbf{L}\frac{1}{r}\frac{dV(r)}{dr}$$

The -1 is from the Thomas precession. With $g = 2$, there is a reduction by a factor of 2, as needed by experiment. One sometimes speaks of a Thomas factor of ½, but more properly it would be $(g-1)/g$. The form of U shows why a subtle and delicate consequence of special relativity, the Thomas precession, can have a gross effect on the energies: The Thomas precession energy is of order $1/c^2$, but so is the spin–orbit energy.

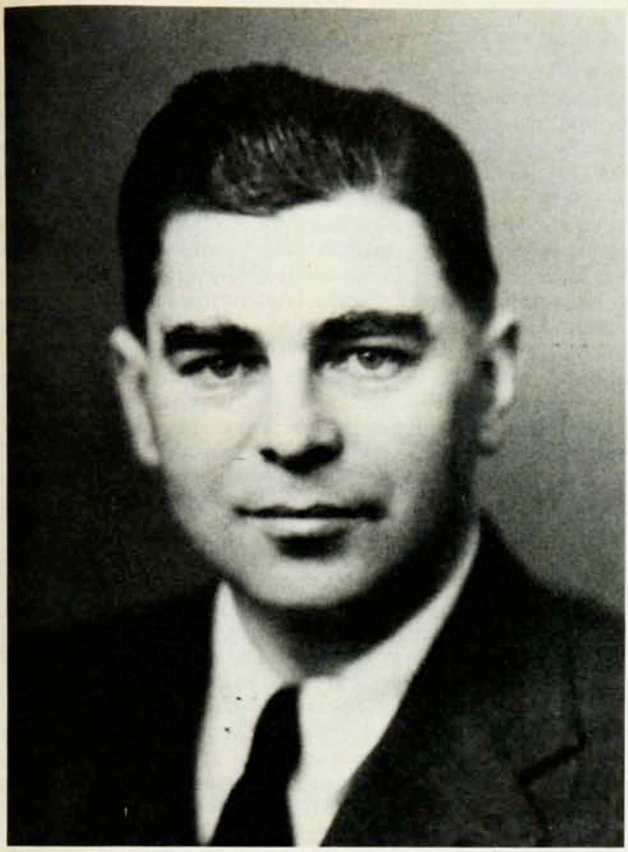
In his longer paper,[8] Thomas describes spin relativistically in terms of a second-rank antisymmetric tensor $S^{\alpha\beta}$ or alternatively an axial four-vector S^α that reduces to a three-vector in the particle's rest frame. Possible covariant equations for spin motion in external fields are greatly restricted by the assumptions that the equation is linear in the spin vector and in the electromagnetic field strengths and that it involves the particle's four-velocity u^α and at most its first time derivative. If the mechanical motion of the particle is described by the Lorentz force equation, the relativistic generalization of the torque equation for a spin magnetic moment in a magnetic field is

$$\frac{dS^\alpha}{d\tau} = \frac{e}{mc}\left[\frac{g}{2}F^{\alpha\beta}S_\beta + \frac{g-2}{2c^2}u^\alpha S_\lambda F^{\lambda\mu}u_\mu\right]$$

Here τ is the proper time, $F^{\alpha\beta}$ is the electromagnetic field tensor, and summation over repeated indices is implied. An equivalent equation, somewhat longer but physically more transparent, can be deduced by considering the time derivative with respect to the *laboratory time* of the *rest-frame* spin vector **s**:

$$\frac{d\mathbf{s}}{dt} = \frac{e}{mc}\mathbf{s}\times\left[\left(\frac{g-2}{2}+\frac{1}{\gamma}\right)\mathbf{B} - \frac{g-2}{2}\frac{\gamma}{\gamma+1}(\boldsymbol{\beta}\cdot\mathbf{B})\boldsymbol{\beta} - \left(\frac{g}{2}-\frac{\gamma}{\gamma+1}\right)\boldsymbol{\beta}\times\mathbf{E}\right]$$

For small velocities (that is, β near 0, γ near 1), one recovers the expected torque on a magnetic moment of $ge\mathbf{s}/2mc$, but at high speeds there are major modifications. The last term (involving **E**) results from the combined spin–orbit interaction and Thomas precession. The first term describes the precession of the magnetic moment in the magnetic field. Note that if the g-factor were exactly 2, the precession would be precisely at the cyclotron frequency $eB/\gamma mc$. The small departure of the electron's and muon's g-factors from 2 causes a precession more rapid by a factor of $(g-2)\gamma/2$; this difference has formed the basis of high-precision measurements of $g-2$ for the electron and muon. The Thomas



Evan James Williams analyzed bremsstrahlung from the electron's point of view, greatly clarifying the physical understanding. Williams initially worked as an experimenter in Ernest Rutherford's lab and then as a theorist specializing in collision phenomena. Born in Wales in 1903 and educated at Manchester and Cambridge, Williams taught at Manchester University, Liverpool University and University College of Wales, in Aberystwyth. During World War II he worked on antisubmarine warfare; he died, back at Aberystwyth, in 1945. (Photo courtesy of Ian Callaghan, Manchester University)

equation forms the basis of the discussion of spin motion in present-day high-energy particle accelerators.[9]

Two comments. The first of the spin-motion equations illustrates explicitly an important aspect of the use of special relativity, the covariance of physical laws. The first postulate, in effect, demands that physical laws be relationships among quantities having the same Lorentz transformation properties. The left-hand side of the spin-motion equation is the proper-time derivative of a four-vector and hence is also a four-vector, because proper time is an invariant quantity. The right-hand side involves two four-vectors and a second-rank four-tensor combined in such a way as to give a four-vector. The mere requirement of covariance or invariance of physical laws is a powerful tool in delimiting the possibilities. Often, as in this example with the assumptions stated above, the result is unique.

The second comment is that 32 years after the fact, Thomas's equations were rediscovered by a group at Princeton who were unaware of the previous publication.[10] Thomas's paper in the *Philosophical Magazine* stands as a *tour de force* with one immediate application, but which had to wait 40 years for broader use (and see the credit go to others).

The Dirac equation

The next example of special relativistic thinking, the Dirac equation,[11] is so celebrated that it might be passed over. Yet it is worth a brief discussion because it shows how an *idée fixe* can have brilliant consequences, even though the driving requirement is later found not to be really necessary. Paul A. M. Dirac's aim in late 1927 was a relativistic generalization of the Schrödinger equation,

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi$$

Because Dirac required the relativistic wavefunction to be interpreted as a probability amplitude, he rejected equations—such as those familiar from electromagnetism and already discussed in quantum mechanics by Oskar Klein, Erwin Schrödinger and others—involving second derivatives with respect to time. The continuity equation emerging from such a theory involves a probability density (if it is interpreted as such) that is not positive definite. For Dirac, this was unacceptable. He therefore returned to the structure of the Schrödinger equation with its first-order time derivative. He then sought an equation that treated energy and momentum on the same (relativistic) footing, with the usual wave mechanical equivalence, according to which $(E, c\mathbf{p})$ is replaced by $(i\hbar\partial/\partial t, -ic\hbar\nabla)$. He was led in a then mysterious, but now familiar, way to an equation involving four-component wavefunctions and 4×4 noncommuting matrices, and first order in space–time derivatives. For physicists the Dirac equation is a thing of beauty and a joy to behold. Further, it contains a seeming miracle, the electron spin, and with the correct g-factor! The demand of linearity in E and $\mathbf{p}$ forced the four-component wavefunction on Dirac. The added degrees of freedom turned out to be the two signs of the spin direction and two signs of the energy (which led to positrons).

The Dirac equation describes quarks and leptons of all sorts, not just electrons. It contains the previously *ad hoc* nonrelativistic spin of Pauli in the appropriate limit (and with its spin–orbit interaction correctly including Thomas precession). Its two signs of the energy emerged later to be the description of particles of spin ½ and their antiparticles. All of these from Dirac's demand for a wave equation first order in the time derivative and so also first order in spatial derivatives—what bounty to be obtained from a requirement that was later found unnecessary!

To clarify briefly the last remark, let me say that particles occur in nature as fermions (with intrinsic angular momenta that are odd half-integral multiples of $\hbar$) or bosons (with intrinsic angular momenta that are integral multiples of $\hbar$). Spin-½ fermions are described by the Dirac equation, but bosons are described by other wave equations. The simplest is the second-order equation rejected by Dirac. When multiparticle descriptions ("second quantization" techniques) were developed, it was found that the "probability density" that was not positive definite is actually a charge density, with its sign depending on the preponderance of particles or antiparticles. One can even develop single-particle descriptions of bosons, analogous to the Dirac equation, based on a relativistic second-order differential equation.

Before leaving Dirac, we must return to the Sommerfeld formula for the energy levels of hydrogenlike atoms. Solution of the Dirac equation for an electron in a fixed Coulomb potential leads to the same energy levels, with the Sommerfeld angular momentum integer k interpreted as $j + \frac{1}{2}$, with j equal to $\frac{1}{2}, \frac{3}{2}, \ldots$. The orbital angular momentum quantum number l does not appear explicitly, only the total angular momentum. Indeed, the idea of separately conserved spin and orbital angular momentum quantum numbers, absent in Sommerfeld's model, but accepted after the work of Pauli and of Uhlenbeck and Goudsmit, is also absent in Dirac's description. Spin *is* an important concept and Sommerfeld did not have it, but its separation from orbital motion is only valid nonrelativistically. Strictly speaking, there is only total angular momentum. For me, this throws some light on the success of Sommerfeld's calculation and the close connection of spin with relativity that Dirac made manifest.

Chronologically, the next major development with special relativity as a key ingredient was quantum field theory. With Dirac's work on quantization of the radiation field pointing the way, a fully general formulation of the quantum theory of fields, with continuous degrees of freedom, was developed at the hands of Werner Heisenberg, Pauli and others. Continuous systems in classical physics, apart from electromagnetism, were known, of course. Euler–Lagrange equations of motion can be derived from the principle of least action, based on a Lagrangian density $\mathscr{L}$ that is a function of the fields and their derivatives. Since quantum mechanics appears most naturally as a generalization of Hamilton's equations of motion for canonically conjugate coordinates, the Hamiltonian played a central role in quantum field theory, with time singled out through, for example, the use of equal-time commutation (or anticommutation) relations for the field operators. Creation and annihilation operators for particles in specific states of motion emerge from normal-mode expansions and the quantization of the simple harmonic oscillator Hamiltonians in the standard way.

While the Hamiltonian is important, the Lagrangian density and its space–time integral, which is the action integral, are extremely valuable constructs. Invariances of the Lagrangian density under various transformations (translations, rotations, internal "rotations" and so forth) lead to conservation laws. The first postulate of relativity imposes the requirement of Lorentz invariance on the action integral and on the Lagrangian density. The Euler–Lagrange equations of motion emerge naturally in covariant form and the condition that $\mathscr{L}$ be a Lorentz scalar restricts mightily the forms of hypothetical interactions among the fields.

The story of the emergence of the Lagrangian, the principle of least action and the use of path integrals as a direct approach to quantum mechanics and quantum field theory is not my task here. It is identified with the name of Richard Feynman above all others, but the idea goes back to Dirac, who in 1933 stressed the superiority of the approach via the Lagrangian because of its manifest invariance with respect to Lorentz transformations.[12] The various approaches to relativistic quantum field theory can be found in standard texts.[13]

In our discussion of random examples of the roles and uses of special relativity in theoretical physics we have come rather far from the practicalities of atomic spectra and Sommerfeld's atom. In part, that is because special relativity permeates the work place of the physicist, if not the layman, from the lowest energies in atomic systems (where the precision is so high that the tiny relativistic effects must be included) to the highest laboratory energies in the giant particle accelerators (where relativistic effects are gross and must enter even the crudest considerations).[14] It is, as I said at the beginning, a pointless exercise to try to count the ways. Instead, I close with examples that illustrate the truth of what Albert Michelson and Edward Morley found: There is no preferred frame of reference.

Frames of reference

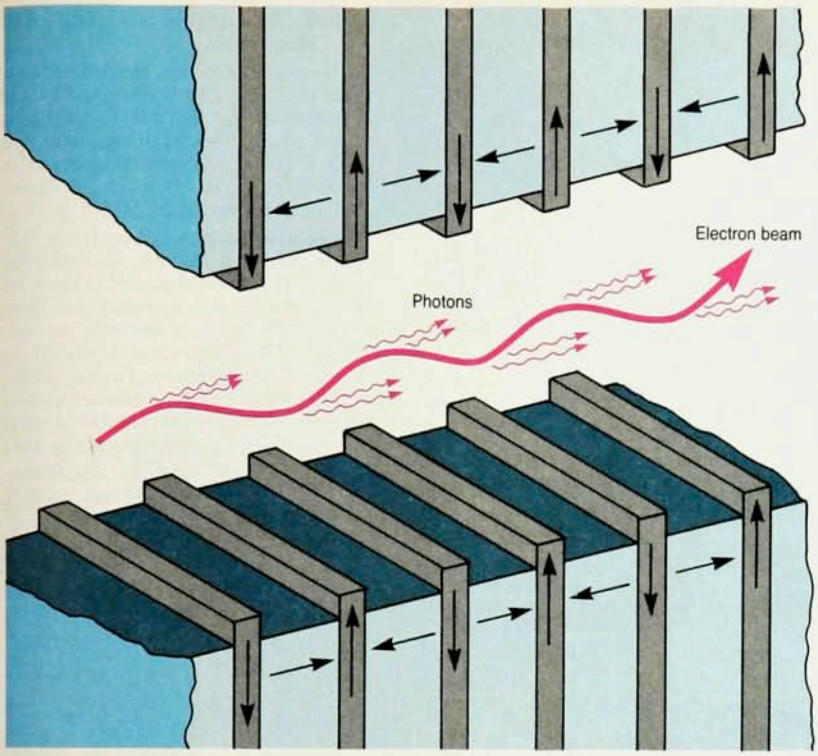
Anyone considering swiftly moving charged particles learns very soon to examine the physical processes in more than one inertial reference frame. Consider bremsstrahlung, that is, the emission of radiation by a light, charged particle on collision with an atomic nucleus. The process was first calculated by Hans Bethe and Walter Heitler, Sommerfeld and others on the basis of second-order perturbation theory in quantum mechanics. But as was shown brilliantly by Carl Friedrich von Weizsäcker and especially Evan James Williams, the physical understanding of bremsstrahlung and many other processes benefits greatly when they are viewed in appropriate inertial frames.

Conventionally, bremsstrahlung is described in the laboratory as the acceleration of a swiftly moving, light, charged particle by the Coulomb field of an essentially stationary heavy nucleus, accompanied by the emission of radiation (see the figure on page 37). In the Weizsäcker–Williams method of virtual quanta,[15] one instead views the process in the rest frame of the incident particle. In this frame the heavy nucleus is incident at high speed upon the stationary light particle. The nucleus's Coulomb field appears, because of the Lorentz transformation, very much like a pulse of electromagnetic radiation, with transverse electric and magnetic fields having a broad frequency spectrum with a number of virtual quanta per unit frequency interval inversely proportional to frequency.[16] These virtual photons, of frequencies ω', are scattered by the stationary light particle into a wide angular distribution—proportional to $\frac{1}{2}(1 + \cos^2\theta')$, at least at low frequencies. If the motion is highly relativistic, these photons, which are predominantly at low frequencies and large angles, are caused by the relativistic Doppler shift to appear in the laboratory at moderate to high frequencies $\gamma\omega'$ and small angles on the order of $1/\gamma$. Bremsstrahlung thus has an alternative description as the scattering of virtual quanta in the rest frame of the incident particle.

There is, in fact, an apparent dilemma, and a neat explanation due to Weizsäcker. The spectrum of virtual quanta of the heavy nucleus contains energies $\hbar\omega$ far above the rest energy of the light particle. Such quanta, when scattered and transformed to the laboratory frame, would have energies on the order of $\gamma\hbar\omega'$, far greater than γmc^2, violating conservation of energy—the largest possible photon energy is $(\gamma - 1)mc^2$. As Weizsäcker showed in detail, however, this violation does not occur, because the scattering of quanta with energies $\hbar\omega'$ greater than mc^2 in the incident particle's rest frame is described by the Compton cross section, not the simple Thomson (recoilless) cross section. The Compton scattering cross section falls off with increasing photon energy, the scattered photons have lower energies because of recoil, and the angular distribution changes. All of these features combine to conserve energy back in the laboratory and give all the details of the less transparent calculations.

A wide variety of inelastic collisions of charged particles can profitably be described by the Weizsäcker–Williams method, whereby the fields of a "projectile" are replaced by an equivalent spectrum of photons interacting with a "target." Usually, but not always, the target is at rest in a reference frame other than the laboratory.

Another example of the usefulness of



Undulator magnet for enhancing the synchrotron radiation from a beam of electrons. Such devices are also the active parts of free-electron lasers. Steel pole pieces are shown in gray; $SmCo_5$ permanent magnets, assembled with alternating polarities, are in blue.

shifting from one reference frame to another is in understanding the frequency of radiation created by electron undulators, devices in which a relativistic electron beam is caused to make very small sideways oscillations by a periodic array of magnets of alternating polarity.[17] The array can be realized by two sets of samarium–cobalt permanent magnets, alternating in polarity, separated by iron pole pieces; the set looks rather like the black and white keys on a piano, with a corresponding keyboard above (see the figure above).[18] The beam passes through the middle and experiences in succession magnetic fields directed upward and downward with spatial wavelength λ_u. The magnetic force causes the beam to undulate from side to side with the same basic period.

Because the electrons are accelerated, they radiate. But at what frequency? To determine it, we first make a Lorentz transformation to the approximate rest frame of the beam. The electrons are almost at rest, so the magnetic force in that frame is negligible, but there is an electric field, equal to $\gamma\boldsymbol{\beta}\times\mathbf{B}$, which causes the electrons to oscillate (because $\mathbf{B}$ is periodic). The oscillating electrons emit dipole radiation in the well-known way. The wavelength is governed by the periodicity of the magnetic field, as seen in this reference frame. Because of the Fitz-Gerald–Lorentz contraction, the magnet structure rushing by has a wavelength λ'_u of λ_u/γ. This is the wavelength of the radiation emitted in the electrons' rest frame. If we now return to the laboratory, we can ask for the wavelength of the radiation there. The relativistic Doppler shift (the Lorentz transformation for light), in the limit of large γ, takes the form

$$\lambda \simeq (\lambda'/2\gamma)(1+\gamma^2\theta^2)$$

where λ' is the wavelength in the moving frame and θ is the angle of the radiation in the laboratory relative to the direction of motion. With the dipole wavelength from the rest frame, the result for the wavelength of the radiation from an undulator is

$$\lambda \simeq (\lambda_u/2\gamma^2)(1+\gamma^2\theta^2)$$

The factor γ^2 means that the radiation can have extremely short wavelengths. For example, an undulator wavelength λ_u of 4 cm is realistic, as is a γ of 3000 (for electrons of 1.5 GeV). The undulator radiation would then have a wavelength of roughly 22 Å (an energy of 560 eV): soft x rays. The motion in the rest frame of the beam is not completely nonrelativistic and sinusoidal. As a result, a number of harmonics of the fundamental are radiated; the relative intensities depend on the details of the undulator. For an undulator of N periods, the radiation peaks have a height proportional to N^2, a fractional width $\Delta\lambda/\lambda$ of order $1/N$ and an angular spread of order $1/\gamma\sqrt{N}$. Undulators provide intense, coherent, tunable sources of soft x rays and vacuum ultraviolet for condensed matter and biological research; they also serve as the "engines" that drive free-electron lasers.

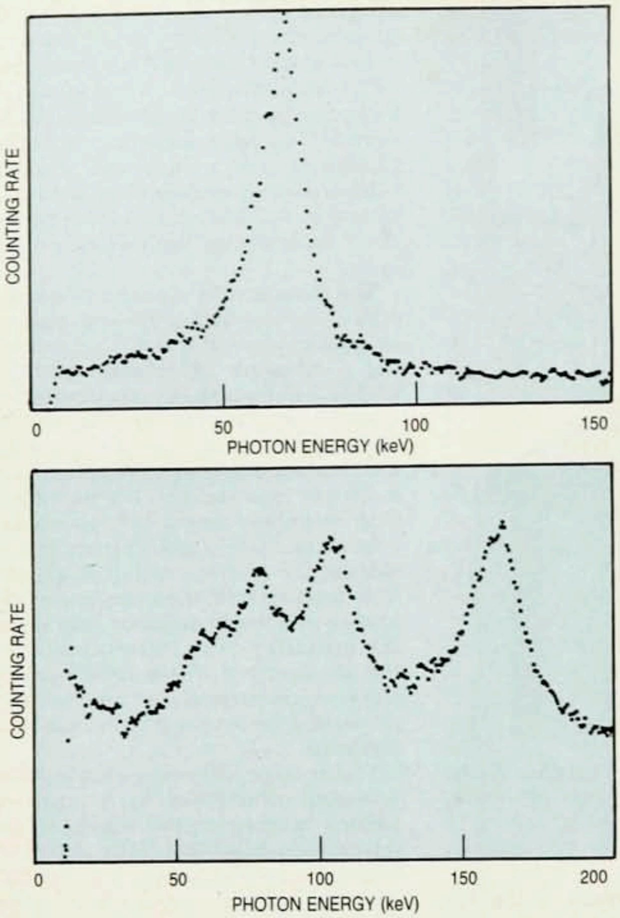
The phenomenon of channeling radiation is my last example of how one can understand essential features by viewing a situation in different inertial frames. If high-energy positrons are incident on a crystal in a direction closely parallel to a crystal plane, they are channeled between the planes like a stream in a straight Alpine valley with mountain ranges on either side. In general, there is motion from side to side as the positron tests the slopes. This transverse motion implies acceleration and hence radiation, just as in the undulator. The frequencies emitted are governed by the basic parameters of the crystal lattice and the relativistic parameter γ of the incident positrons.

The average transverse electrostatic potential experienced by a positron passing between crystal planes is approximately quadratic. (The period of transverse oscillation is long compared with the transit time for one lattice spacing. The longitudinal periodicity is thus washed out.) The "spring constant" $V''(0)$ for the effective oscillator can be estimated by dimensional arguments to be roughly 13.6 eV$/a^2$, where a is the distance between lattice planes. The transverse electric field in the laboratory, $\mathbf{E}_\perp$, is $V''(0)\mathbf{x}_\perp$. If the positron were nearly at rest in the laboratory, its transverse frequency of oscillation ω_0 would simply be $(V''(0)/m_e)^{1/2}$, corresponding to optical photon energies of a few electron volts.

Let us go to the moving coordinate frame in which the positron is almost at rest, the frame moving with the velocity of the incident beam. There the positron experiences the transformed electrostatic force of the lattice. This has electric and magnetic components, but since the positron's motion is nonrelativistic in this frame, only the electric force is important. The transformation law of the transverse electrostatic field is

$$\begin{aligned}\mathbf{E}'_\perp &= \gamma\mathbf{E}_\perp\\ &= \gamma V''(0)\mathbf{x}_\perp\\ &= \gamma V''(0)\mathbf{x}'_\perp\end{aligned}$$

(The transverse coordinates are, of course, unchanged by the Lorentz transformation.) The "spring" is thus stiffer by a factor of γ in the positron's average rest frame, and the frequency



Channeling radiation spectra for 55-MeV positrons (**a**) and electrons (**b**).[21] The particle beams are parallel to the (110) planes of diamond crystals. Note that the positron spectrum exhibits a single peak, while the electron spectrum is highly structured.

of its transverse motion is $\omega_0\sqrt{\gamma}$. The dipole radiation from the oscillation appears in the laboratory in the usual narrow forward cone with angular width of order $1/\gamma$ and with a frequency ω given approximately by $2\gamma^{3/2}\omega_0/(1+\gamma^2\theta^2)$. For 50-MeV positrons, γ is around 100 and the photon energies are 2000 times optical, or tens of keV.[19] For 10-GeV positrons, the energies are 10^7 times optical—around 50 MeV![20]

To the extent that the potential is truly quadratic, even a quantum mechanical treatment gives only the fundamental line. Observations always show a strong peak at the fundamental, sometimes with smaller peaks near multiples of the basic energy. For high-energy electrons the emission occurs in the same general energy range, but the electron spectra can exhibit a number of prominent peaks.[19,21] (See the graphs above.) The difference arises because the electrons experience a funnel-shaped electrostatic interaction, not a simple harmonic oscillator potential. The frequency of classical transverse motion then depends on the amplitude, and for a given amplitude the periodic motion corresponds to a superposition of many harmonics of the basic frequency. (In quantum language, the energy levels of transverse motion are not equally spaced and transitions are not restricted to those between adjacent levels.)

These examples are but illustrative of the value of viewing physical phenomena in different inertial frames. The absence of a preferred frame of reference, established by the Michelson–Morley experiment 100 years ago and bothersome to many at the time, has been turned with Einstein's guidance into a very positive virtue. I have tried to document how a few practitioners used special relativity to good effect in its early days, and how it still gets used to simplify and clarify our understanding.[22] Equivalent inertial frames are everywhere. Use them for profit and enjoyment!

References

1. W. Pauli, *Relativitätstheorie* (*Enzyklopädie der Mathematischen Wissenschaften*, vol. 19), Teubner, Leipzig (1921).
2. A. Pais, *Inward Bound*, Oxford U. P., New York (1986).
3. For a succinct account, see A. Pais, *Subtle is the Lord: The Science and Life of Albert Einstein*, Oxford U. P., New York (1982), ch. 7. The details are described by A. I. Miller, in *Some Strangeness in the Proportion*, H. Woolf, ed., Addison–Wesley, Reading, Mass. (1980), p. 66.
4. Sommerfeld's original paper is not easily accessible. In English, an authoritative source is the translation of the third edition of his book *Atombau und Spektrallinien*: A. Sommerfeld, *Atomic Structure and Spectral Lines*, H. L. Brose, transl., Methuen, London (1923), ch. VIII, sect. 2 and mathematical note 16; see also sect. 7. A number of texts from the 1930s present his derivation.
5. A. Pais, *Subtle is the Lord: The Science and Life of Albert Einstein*, Oxford U. P., New York (1982), p. 408.
6. E. Segrè, *From X-Rays to Quarks*, Freeman, San Francisco (1980), ch. 8. A. Pais, *Inward Bound*, Oxford U. P., New York (1986), ch. 13.
7. L. H. Thomas, Nature **117**, 514 (1926).
8. L. H. Thomas, Philos. Mag. **3**, 1 (1927).
9. B. W. Montague, Physics Reports **113**, November 1984, page 1.
10. V. Bargmann, L. Michel, V. L. Telegdi, Phys, Rev. Lett. **2**, 435 (1959).
11. P. A. M. Dirac, Proc. R. Soc. London, Ser. A **117**, 610 (1928).
12. P. A. M. Dirac, Phys. Z. Sowjetunion **3**, 64 (1933); reprinted in *Selected Papers on Quantum Electrodynamics*, J. Schwinger, ed., Dover, New York (1958), p. 312.
13. See, for example, C. Itzykson, J.-B. Zuber, *Quantum Field Theory*, McGraw–Hill, New York (1980).
14. W. K. H. Panofsky and E. M. Purcell address aspects of relativistic engineering in *Some Strangeness in the Proportion*, H. Woolf, ed., Addison–Wesley, Reading, Mass. (1980), p. 66.
15. C. F. von Weizsäcker, Z. Phys. **88**, 612 (1934). E. J. Williams, Det. Kgl. Danske Viden. Selskab., Math-fys. Medd. **13**(4) (1935). See also E. J. Williams, Proc. R. Soc. London **139**, 163 (1933).
16. The spectrum and some of the applications are discussed in a number of texts. One is J. D. Jackson, *Classical Electrodynamics*, 2nd ed., Wiley, New York (1975), sect. 15.4.
17. H. Motz, J. Appl. Phys. **22**, 527 (1951).
18. D. Attwood, K. Halbach, H.-J. Kim, Science **228**, 1265 (1985).
19. M. J. Alguard *et al.*, Phys. Rev. Lett. **42**, 1148 (1979). R. L. Swent *et al.*, Phys. Rev. Lett. **43**, 1723 (1979).
20. J. Bak *et al.*, Nucl. Phys. B **254**, 491 (1985).
21. H. Park *et al.*, J. Appl. Phys. **55**, 358 (1984).
22. The effectiveness of the use of different inertial frames is illustrated masterfully in E. M. Purcell, *Electricity and Magnetism* (*Berkeley Physics Course*, vol. 2), 2nd ed., McGraw–Hill, New York (1985). □