# SPACE, TIME AND ELEMENTARY INTERACTIONS IN RELATIVITY

To express in a unified formalism all the interactions of matter, ranging from the elementary particles to astronomical bodies, we should begin with a simple study of the concepts of Einstein's space—time.

MENDEL SACHS

A REASONABLE COURSE of investigation in fundamental physics should fully exploit Einstein's approach to the meaning of space-time; it would proceed by exploring, in a unified way, the predictions of generalized equations of general relativity that would encompass in one formalism all of the domains of interaction—from fermis to light years.

#### SEARCH FOR SIMPLICITY

When the theory of relativity was first proposed and had attained some success in its initial predictions, a rumor started to circulate that throughout the world only a very small number of men actually understood this theory. The general impression was thereby created that this was so profound an approach to the laws of nature that most ordinary scientists should not even attempt to understand the foundations of this theory. But this impression in the scientific community was incompatible with one of Einstein's prime motivations in his investigation. For he was very strongly guided in this study, and indeed throughout his scientific career, by the heuristic value of simplicity.

Around the same time that relativity theory started to develop, the other revolution in 20th-century physics the quantum theory—was discovered to reveal a formalism that explained

atomic phenomena. This theory eventually developed into its present form of a nondeterministic theory of measurement. It is my contention that the quantum theory is much more difficult to understand from the conceptual point of view than is the theory of relativity. It is partly for this reason that some of the original founders of the quantum theory (in its old form), such as Max Planck and Einstein, and later Erwin Schrödinger, could not accept the notions that were proposed by the Copenhagen school to underlie the fundamental description of matter. On the other hand, the mathematical expression of the new quantum theory -its equations and the rules for relating their solutions to observables-is far more simple than is the general mathematical expression of the theory of relativity.

We see here that a distinction must be made between conceptual simplicity and mathematical simplicity. When Einstein expressed the belief that "God may be subtle but He is not malicious," I believe that he was referring to his faith in the simplicity of the conceptual content of the laws of nature. On the other hand, if man has thus far been unable to formulate conceptually simple, though subtle, natural laws in equally simple language, this is not God's fault! It appears to me that this inability is rather due to the relatively primitive stage of

man's intelligence. I should like, then, to concentrate here on the conceptual content of relativity theory, as I see it and without writing down a single equation, to see what this approach implies with regard to the properties of nature.

## A REVOLUTIONARY CONCEPT

It is commonly admitted that the theory of relativity brought a revolution in ideas to the 20th century. Nevertheless, there seems to be some controversy among physicists on precisely what it is that is revolutionary



Mendel Sachs, who took his bachelor's, master's and doctor's degrees in physics at UCLA, is now a theoretical physicist at the State University of New York at Buffalo. A member of the editorial board of the International Journal of Theoretical Physics, he has worked recently on the physical implications in the elementary-particle domain of generalizations of general-relativity theory.

here. My contention is that Einstein's original intention in the change that he instigated had to do with an entirely new view of space and time. That is, except for the argumentation of some of the preceding philosophers, Einstein gave a view of space-time that was different from all the previous views in physical theories. This change, in turn, implied an alteration of the form of the laws of nature to a more generalized expression that could be tested by experimentation.

# Pre-Einsteinian relativity

What, then, is the actual difference in the view of space and time between Einstein's approach and that of his predecessors (and conservative contemporaries)? Consider first the earlier theories. In all the previous interpretations of the physical universe, it was assumed that space and time are just there, once and for all, independent of any matter content. Matter is then supposed to be in the preexisting space and time, as a child might be on a jungle gym in the playground. One then describes how matter interacts with matter in terms of its location in the fixed space-time. In philosophical terms, this description may be restated as the view that space-time has objective significance, that it is a thing in itself.

It is important at this point to note that the relativity of coördinates within an underlying fixed space-time in the description of the natural laws is not a new idea and that the idea of relativity in this sense is really not such a revolutionary concept. Consider, for example, the physical laws before the time of Galileo. It would have been claimed by most in that period that "up" is "up" to anyone-independent

## SPEED OF LIGHT

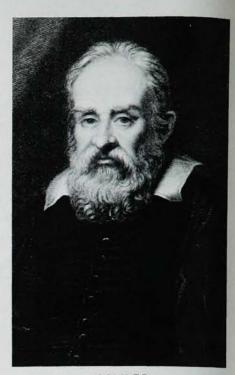
In my opinion it is not necessary to assert the universality of the speed of light as a separate postulate, as is conventional in special-relativity The assertion appears rather to follow as a logical conclusion of the single underlying postulate of this theory—the principle of relativity (page 53)—as soon as the coördinates of relatively moving systems are defined in a way that assigns the temporal as well as spatial parameters to the language of one observer relative to the others who are comparing their respective deductions about the form of the laws of nature.

of his position or state of motion relative to anyone else. Similarly it would have been claimed that "before," "present" and "after" have absolute meaning in the sense of being independent of the condition of motion or the relative position of any observer. On the other hand, it would have been generally agreed that in the description of a physical law the location of a point in a two-dimensional plane is only significant in relation to some other points in the plane-that is, physical laws depend on distances between things. For example, everyone experienced the diminishing heat effect of a flame as he correspondingly increased his distance from the location of the fire.

Then a time came, with the appearance of Copernicus and Galileo, when the natural philosopher was forced to incorporate the "up" and the "down" with the surface coördinates to yield a three-dimensional space where, for example, "up" to one observer might be "down," "sideways" or "oblique" to another observer of the same phenomena. Finally, more than 300 years after Galileo, Einstein recognized that the time coördinate must also necessarily be incorporated with the three spatial coödinates in one relativistic space-time to describe the laws of nature correctly. This discovery was made when Einstein recognized that the laws of electrodynamics, in the form that they were originally discovered by Michael Faraday and James Clerk Maxwell, have the same form to all observers, independent of their states of motion relative to each other, only if it can be assumed that the time coordinate is included with the spatial coordinates in the relativistic sense that to describe an observation, all four coördinates must be specified relative only to the space-time point of a given observer. At this stage, the theory of special relativity was born.

#### Unification of space-time

Incorporation of the time coördinate with the spatial coördinates implied some interesting consequences, not predicted previously. First, to be able to convert the time coördinate of one observer to a combination of time and space coördinates of another (who is in motion relative to the first) one must express time and length in the same units—in all frames of reference. Thus the time coördinate had to be multiplied by a *universal constant* with the dimension of length divided by



GALILEO

time-a constant speed. Indeed, the identification of the laws of electrodynamics with this number revealed it to be the speed of propagation of the electromagnetic interaction between matter and matter in a vacuum-numerically determined from the speed of light in a vacuum. This determination of the universal speed c then fixed it for all other applications (see box on this page). A further implication here was that the speed of light in a vacuum is independent of the speed of its source. That is, according to this theory, if light is emitted from a moving train or from a fixed position next to the tracks, a fixed observer will measure the same speed in each

Another new consequence of relativity theory was that no longer could we speak of the absolute simultaneity of events. If two events are simultaneous to one observer, they are not generally so to a different observer in motion relative to the first one. A third interesting consequence of special relativity came from the modification that results when we pass from classical mechanics of point particles (where time is absolute) to relativistic mechanics of point particles (where the time, along with the spatial coordinates, is relative to the point of observation). It turns out in this generalization that although the inertial mass of an interacting body is constant when the body is at rest, it increases in a definite way as the body moves relative to a fixed observer. Further, when at rest, the body has an intrinsic energy that depends on the product of its rest mass and the square of the speed of light—an explosively large amount of energy that was not suspected until relativity theory appeared! These and other implications of special-relativity theory in regard to wave motion, for example, that were not predicted by the preceding theories were impressively substantiated by experimental facts.

Thus we see that the rules of special-relativity theory refuted the previous laws of mechanics by requiring that the time coördinate must be incorporated with the three spatial coordinates as a fourth parameter, to be specified only relative to the observer. But this change was not too much of a revolution! It was rather a natural extension to a more general way of expressing the laws of nature with a space-time coöordinate system. In this case, then, what was the real revolution that came with Einstein's theory? It was the abandonment of the idea that the space-time coordinate system has objective significance as a separate physical entity. Instead of this idea, relatively theory implies that the space and time coördinates are only the elements of a language that is used by an observer to describe his environment. However, the conventionalist view is not fully adopted since it is further asserted with this theory that the relation between the points of the space-time language of any observer is in fact a representation of the intrinsic interaction within the matter distribution that comprises the physical system. It then follows that if there should be no matter in the universe, there would be no space-time to talk about! It implies that if the matter distribution should be variable, then the relation between the points of the space-time (that is, the geometry) that is used to describe the environment of any point would be correspondingly variable. This view takes space-time as a passive entity that is used to describe nature-perhaps for want of a better language!

# THE PRINCIPLE OF RELATIVITY

The fundamental starting point of relativity theory—the principle of relativity—asserts that, if the laws of nature that are substantiated by one observer's data are indeed bona fide natural laws, their form cannot change when they are deduced by any other observer

"... the theory of relativity introduced a revolutionary concept into physics by relegating space and time to the subjective role of the elements of a language that one observer or another may use to describe the natural laws."

who may be at any other space-time point and in arbitrary relative motion with reference to the first observer. Let us examine this assertion a little closer. It means that after the natural laws have been deduced and substantiated in one space-time frame of reference, with the language (x,y,z,t) of one observer (who need not be a human or his equipment; it may also be an electron!) then in any other frame of reference, with the language (x'y',z',t') of the different observer who is in arbitrary relative motion compared with the first observer, the derived natural laws must be identical in form. (See box on page 52).

Of course, the observer who uses the language (x,y,z,t) and the other observers who use the languages (x'',y'',z'',t''), (x'',y'',z'',t''), . . . would never be able to check that their respective laws of nature for the

same sorts of phenomena are actually the same until they learn to translate from one language to the other. The collection of translations of all words (space-time points) from one coordinate frame to the others, which may be in arbitrary relative motion, is a continuous transformation group; it is called the "Einstein group." The representations of the group provide a powerful mathematical device for abstracting many of the implications of relativity theory that are not at all obvious at first glance. Some of these implications in regard to general relativity will be discussed later.

We see, then, that the theory of relativity introduced a revolutionary concept into physics by relegating space and time to the subjective role of the elements of a language that one observer or another may use to describe the natural laws. Use of the theory is



EINSTEIN

subject to the restriction that the form of these laws must be unchanged under the cöordinate transformations between any space-time frames of reference that are in arbitrary relative motion.

#### ELEMENTARITY OF INTERACTION

If space and time are indeed to be considered in the subjective role that was discussed above, and if at the same time we adopt the realist view that there are fundamental objective elements underlying the laws of nature, then what does relativity theory imply these to be? That is to say, in the earlier theories, and even in the contemporary elementary-particle theory, the world is made out of things -indestructible bits of matter that move along space-time trajectories. The quantum-mechanical approach is of course not quite as deterministic as classical mechanics, because the classical mechanics makes precise predictions about the trajectories of interacting things, and the quantum theory must resort only to probability statements about the states of motion of these things. Nevertheless, it is the bit of matter, with its own space-time trajectory, that is, in both of these theories, the objective reality out of which the world is constructed. Each of these mechanical theories is indeed a different version of an atomistic approach to the natural laws.

On the other hand, when one follows the axiomatic basis of relativity theory to its logical extreme, it appears to me that it is not the free bit of matter that is the fundamental building block with which to construct a theory of matter. It is rather a relation that forms the basic entity here. With relativity theory, one must take the "observer-observed" relation as a fundamental starting point. That is to say, the "observer" is here an entity with meaning only in relation to the "observed," and vice versa. Also, in contrast with the quantum theory, "observer" has no anthropomorphic denotation. It refers only to an intrinsic component of a fundamental interaction. It might refer to a component of a physical system that could be identified with a star, a man or a proton! On the other hand, the "observer" and the "observed" in classical and quantum mechanics do indeed have meaning as things in themselves. Again, in contrast to the quantum theory, the relativistic theory requires a

"Relativistic theory requires a theoretical structure in which it makes no difference to the predictions of the theory which part of the physical system is identified with the 'observer' and which part is called 'observed."

theoretical structure in which it makes no difference to the predictions of the theory which part of the physical system is identified with "observer" and which part is called "observed." These names are merely chosen for reasons of convenience to describe the components of a single interaction in one space-time—without actual parts!

To compare the predictions of this theory with certain (but not all) experiments, one must be able to examine this one relation in the asymptotic region, where there is sufficiently weak intrinsic coupling to reveal an apparent uncoupling into separated parts. It is important, however, that with the assumption about the elementarity of this relation (which may be called "elementary interaction") there can never in principle be any actual separation, no matter how closely one may approach this apparent situation in the theoretical description. Such a conclusion about the intrinsic inseparability of the elementary interaction does not refute any experimental evidence because it is, in principle, impossible to observe a completely uncoupled component of a physical system. The very act of measurement automatically couples an apparatus (not necessarily macroscopic!) to the "observed." But, independent of approximation, it is still the one relation-"elementary interaction"-that is the fundamental entity in a fully exploited theory of relativity. Such an entity is of course entirely objective as its fundamental description is independent of the interchange of the names for the subject (the "observer component") and the object (the "observed component").

#### A metaphor: a dog and a flea

This idea can be illustrated with the following simple allegory. Consider a world that consists of a dog, a flea, the sky and the ground. Let us say that the dog's life span is of the order of 15 years and that of the flea is about 5 hours. Suppose now that the flea

lives on the dog's back, and let us inquire about the flea's and the dog's respective accounts of the world according to the atomistic and relativistic approaches. According to the atomistic view, the flea will think that he lives in a forest of tall soft trees, on warm, undulating ground. He would often recall the painful experience that occurs every few years of his life, when a major earthquake throws him up into the sky-only to land again in a strange land, but still surrounded by tall soft trees on warm moving ground. The flea would also think about himself-that he is strong, handsome, clever, etc. . . (this is called "selfenergy" in modern physics). The dog's idea of the world, on the other hand, would be quite different. He would be aware of the cool still ground. But he would also worry and be annoyed by a terrible itch on his body that recurs every few minutes, even though he has thoroughly shaken his body. The dog would also think about himself-that he is a tall handsome dog with a beautiful long tail, and that he is clever, modest, etc. . . .

After considering both stories, it would be impossible for us to decide which is the *true* description of the world. The reason, of course, is that neither is the complete story! Each of these accounts presents a subjective view; one belongs to the flea and the other to the dog.

Suppose now that we consider the relativistic view in which it is the relation that is the elementary entity. In this case, the flea would proceed as before to think about the warm undulating ground with the tall soft trees and the periodic earthquakes every few years. However, instead of going into self-attributes, the flea would make a careful study of his environment (the dog). If he is a sufficiently perceptive investigator, the flea may be able to deduce precisely how the dog is Similarly, reacting to his presence. the dog would proceed to think about his environment as he did before. But he would not try to relate his self-attributes to an objective description of the world; instead he would learn how his environment (the flea) is reacting to him. We see, then, that with this approach the flea and the dog would have identical descriptions of the world-even though the languages with which they express these views are entirely different. With this relativistic approach, the world would be described with a single relation, dogflea, without parts! The description is now entirely objective because it is indeed independent of which component of the coupled system is expressing the laws of nature.

#### THE UNIFIED FIELD CONCEPT

The basic elements of the language of a fully exploited theory of relativity are continuous field variables. These are continuously mapped functions of the space and time coordinates. This feature is a consequence of the assumption that the valid laws of nature are invariant in form under the transformations of the Einstein group (the group of continuous transformations among the space-time frames of reference that are in arbitrary motion relative to each other). Einstein showed. for example, that even with the idea of a finite propagation time for the interaction between distant bits of matter, it is not possible to consider this matter in discrete quantities within relativity theory. It would not be possible in this case, for example, to define the conserved quantities of the system, such as energy and momentum. The approach of relativity theory then necessitated an expression in terms of field equations, where the field variables relate to densities. In the final analysis, to compare the predictions of the theory with the experimental facts, it becomes necessary to integrate certain prescribed functions of these field variables over all of the coordinates in which they are mapped. Thus we see that the basic elements of the language in relativity theory are not the space and time coördinates themselves but rather a certain set of functions that are continuously mapped in the spacetime coördinate system.

## Faraday's view of gravitation

The field concept was originally introduced to physics by Faraday about 50 years before the discovery of relativity theory. Before Faraday there was little (recorded) doubt in any physicist's mind that the fundamental description of the material world must necessarily be in terms of bits of matter that act on each other at a distance, for example in the way that Newton envisioned the force of the sun on Earth and the other planets. On the other hand, Faraday recognized that Newton's gravitational force need not be viewed in terms of this model. He rather saw the effect of the sun on the planets in terms of a continuous field of force-the continuous function 1/R2 of the distance R from the center of the sun to any exterior point of obser-Thus, instead of assigning special meaning to the separate spatial points that locate the sun and the planets, the field approach of Faraday considers one space, and a special continuous function of its coördinates predicts how a test particle would move should it be placed at any spatial location. That is to say, Faraday's approach took the potential field of force (a continuous entity) to replace the bit of matter as a fundamental construct. The conflict of these two approaches is indeed as old as the study of physics itself; it is the conflict between discreteness and continuity in the fundamental description of matter.

## An interpretation of the field

Faraday also believed that all the different manifestations of the influence that matter exerts on matter are derivable from a single unified-field description. Of course he was unable to move planets around in order to test

this hypothesis in regard to the motion of the heavenly bodies. He did initiate his investigation of this idea in his studies of electricity and magnetism, physical phenomena that were previously thought to be unrelated. After studying the topological features of the electric and magnetic fields, by respectively observing the electrostatic potential in an electrolyte and the pattern of iron filings in the vicinity of a magnet, Faraday proceeded to investigate whether or not these two fields were unified, in a single electromagnetic force field. He saw that this was indeed the case by observing that, when an electrical conducting wire is moved across a magnetic field of force, an electric potential is generated thereby causing an electric current to flow. Similarly when a current is caused to flow through a wire he observed that a compass needle will polarize in a plane perpendicular to the direction of the electric-current

An important feature of Faraday's view of the field as a continuously mapped potential of force is that this entity would lose all meaning as such should there not be any test charge as a probe. That is to say, because the electromagnetic field was interpreted by Faraday solely as a preexisting cause for an observable effect -the motion of a charged test particle -and because cause and effect are logically inseparable, the electromagnetic field of force in Faraday's view was meaningful only when a test



FARADAY



MAXWELL

"... if space and time are only to serve the passive function of providing a language for the observer, why should the inertial frame of reference in particular be singled out?"

charge was taken to exist at the same time. However, the structure of Faraday's electromagnetic field of force was taken to be uninfluenced by the presence of a test charge in the physical system. This field was taken, rather, as a fundamental representation of charged matter outside of the test charge. Indeed the electromagnetic field of force was assumed by Faraday to be the thing in itself that replaces the bit of matter of "particle" theories as the basic entity from which the elementary description must be built.

Thus, in contrast with the "particle" theories (for example, Newton's theory of gravitation) the fundamental thing in itself in Faraday's field approach played the role of the cause in a cause-effect relation—a relation that is logically inseparable! Atomistic theories, on the other hand, are based on the "free particle" (say, the sun) as the fundamental reality—an entity that has meaning with or without other particles in the system. Later we will see that the assumption of the test



RIEMANN

charge, as uncoupled from the force field, must be removed in the exact form of a relativistic theory, simply because of the elementarity of the interaction that follows from a logical analysis of this theory.

In some of his later experimental studies, Faraday was unfortunately unsuccessful in extending his earlier results to an electromagnetic-gravitational force field. However, in addition to the lack of experimental verification, there were still some basic theoretical questions that had to be answered about such a unification, even at that time. For example, if electromagnetic and gravitational phenomena are in fact manifestations of the same force field, why is it that gravitational forces are only attractive but electromagnetic forces can be. either attractive or repulsive? Later I will show how the answer may lie in the necessary incorporation of the electromagnetic and gravitational force fields with fields that relate to the inertia of matter-fields that were not suspected to exist in 19th-century physics!

## Maxwell's field equations

It was the genius of Maxwell that created the exact mathematical description for Faraday's electromagnetic force field. After formulating "Maxwell's equations" he discovered that not only did some of their solutions relate to the predictions of electromagnetic phenomena that were already known, but that there were other solutions of these equations, describing the faraway effect of oscillating charged matter, that precisely describe the properties of light. Thus Maxwell found that all optical phenomena could be described with the "radiation" solutions of the electromagnetic equations, thereby unifying the optical with the electromagnetic manifestations of interacting matter in a single field theory. This discovery settled a very old dispute (at least temporarily!) about the corpuscular or continuous nature of light. Of course, other manifestations of electromagnetic radiation were discovered not long after this—x radiation, radio waves, gamma radiation, etc.

The salient feature of Maxwell's equations that Einstein discovered is that they are form invariant under a special class of continuous transformations among the space-time coordinates (x,y,z,t) of one frame of reference and the coördinates (x',y',z',t')of other frames of references that are in relative motion. These are the geometrical transformations that leave the interval between any two such sets of coördinates,  $[c^2\Delta t^2 - (\Delta x^2 +$  $\Delta y^2 + \Delta z^2$ ]<sup>1/2</sup>, unchanged. Here c is the constant speed that appears in Maxwell's equations. The particular form of this interval in the fourdimensional space is a characteristic feature of special-relativity theory. If one should choose to keep the time coordinate fixed, as a special case, (by taking  $\Delta t = 0$ ) these transformations would reduce to the set of coordinate changes in a three-dimensional space characterizes the symmetry Galileo discovered to underlie the laws of classical mechanics. But, according to the structure of Maxwell's equations, one may not generally fix the time coördinate, as was done in classical mechanics, if the particular laws of nature that are embodied in Maxwell's equations are to be the same to any set of observers who are in relative motion. It then follows, from this analysis of the laws of electromagnetic fields, that the laws of mechanics that obey this more general transformationgroup invariance (the theory of special relativity) are necessarily different from the laws of classical mechanics. However, it turns out that if the relative speed between interacting bits of matter is small compared with the universal speed c (the speed of light in a vacuum) the two laws of mechanics are practically the same. The differences in the predictions of the respective formalisms are too small to have been detected by the type of experimentation that preceded the 20th century.

#### UNIVERSAL INTERACTION

The Maxwell formulation was the first set of field equations satisfying the invariance requirement of the theory of special relativity to be discovered. Einstein also noticed that special relativity, although it incorporates the time coördinate with the spatial coordinates, relates only to inertial frames of reference.

These are space-time frames that are moving with the special feature of constant rectilinear speed (or are at rest) relative to each other. The theory of special relativity was not acceptable as the most general description of matter because the inertial frame refers only to an extrapolation from observations that actually entail noninertial frames of reference. The reason is that the inertial frame refers only to a description in which there are no forces involved and therefore where there is no energy and momentum transfer between interacting matter. On the other hand, any measurement (whether it involves the coupling of only microscopic matter, macroscopic matter, or a combination of these) necessarily entails a transfer of energy and momentum, if something is to be recorded about the "observed." Secondly, Einstein argued that if space and time are only to serve the passive function of providing a language for the observer, why should the inertial frame of reference in particular be singled out? He then concluded that to exploit this theory fully, the principle of relativity must necessarily imply that the laws of nature are invariant in form under the transformations between space-time frames of reference that are in arbitrary relative motion.

Extension of his theory to noninertial frames of reference then led to the theory of general relativity, with special relativity serving only as a particular limiting case. I should say here that in view of the elementarity of the interaction that follows from the conceptual starting point of relativity theory, the actual limit of special relativity cannot in principle be reached, even though it can be approached arbitrarily closely. this view, then, the success of the mathematical formulations in modern physics that are consistent with special relativity is not more than an indication, in specific applications, of how good special relativity is as a mathematical approximation for field equations that are normally expressed in general relativity.

Had Einstein stopped with special relativity, although it was eminently successful in explaining and predicting experimental facts not previously understood, not too much of a revolution would have occurred. For, as in the previous "classical" theories, the relation between points in space-time in special relativity is the same everywhere, and it is independent of the matter content of the physical system that is described. Fully to exploit the idea that the space and time coordinates serve only as a language used to represent the matter content of a physical system, it becomes necessary to extend from the flat (Euclidean) space-time geometry of special relativity to a curved (non-Euclidean) space-time of general relativity. The variable curvature of the geometrical system in the general description is indeed a function of the variation of the mutual interaction within the physical system. According to this view, if the matter content of a system was depleted the curvature of space-time would correspondingly diminish. In the limit, where the system is emptied of matter (corresponding to no mutual interaction of matter), the curvature would be zero, thereby yielding a Euclidean geometrical system and special relativity. The particular non-Euclidiean geometry that has this property of asymptotically approaching a Euclidean geometry is that discovered by Georg Riemann. Einstein looked for a relationship between the field of mutual interaction of the matter content of a physical system and the field properties of a Riemannian space-time.

The Einstein field equations are a special representation of a relationship between the metrical field (the function that prescribes the relation between points in a Riemannian fourspace) and the mutual interaction of matter fields. In particular, the left-hand side of Einstein's equations are certain nonlinear differential forms in the metric tensor, while the right-hand side depends on the energy-momentum tensor for the matter content of the physical system. Thus, if one is given the energy-momentum tensor for the system, the solutions of these

equations thereby yield the corresponding metric tensor, which in turn prescribes the variable relation between the points of space-time. If one now takes seriously the contention that the metrical field is a representation of the matter that is being described, according to the restraint of Einstein's equations, then it appears to me to follow that these field equations must be considered as "if-andonly-if" relations, rather than "if-then" relations. With this conclusion, the solutions of these equations that correspond to empty space must be rejected as physically unacceptable. Similarly the inversion of these equations and the insertion of the metric tensor for a flat space must yield matter-field solutions that are physically unacceptable. It follows then that space-time can only be described as curved, in any realistic situation, even though it can come arbitrarily close to flatness. This conclusion is compatible with the one that was reached earlier indicating that the limit of special relativity may be approached arbitrarily closely (indeed the empirical facts require this to be so) but that the limit cannot actually be reached within the framework of this theory. Some of the physical implications of this result will be discussed below in connection with the concept of inertial mass.

# Equations of motion

One of the very important features of Einstein's field equations is their containment of the equations of motion of interacting masses. In this respect they contrast with Maxwell's equations for electromagnetism. Indeed, to predict the equations of motion of interacting charged bodies, Lorentz had to adjoin further equations to the Maxwell formalism; these equations then define the electric and magnetic forces (in terms of the coupling of matter source variables to the electromagnetic field-intensity solutions of Maxwell's equations). In the

"In principle, the same set of field equations that describe the domain of elementary-particle physics must also describe the domain of laboratory dimensions and the domain of astronomical dimensions."



MACH

gravitational problem, however, when one recognizes the experimental fact that the inertial mass of a body in nonuniform motion (for example, centrifugal motion) is equal to the mass that is acted on by a gravitational field, then one sees that the effects on bodies that result from the gravitational forces exerted by other bodies are already incorporated in Einstein's metrical field equations. The relation between the points in space-time, when there is nonzero curvature, predicts the gravitational force between massive bodies-without the need to add further equations of motion. Einstein concluded from this feature of general relativity that perhaps not only the gravitational forces but all other forces between matter and matter may be different manifestations of an even more generalized geometry for spacetime. This conclusion, of course, supported Faraday's intuitive feelings about a unified field theory. equivalent statement about Einstein's conclusion regarding the role of geometry in physics follows from an interpretation of his equations as "if-andonly-if" relations. For it would then follow that a generalized geometry of space-time could be viewed as a manifestation of all possible forces between matter and matter-a manifestation of a universal interaction.

The first test of Einstein's equations was to see if they would predict all of the effects that Newton had already successfully predicted about gravitational forces. Einstein found that, under the special approximating cir-

cumstances that are accurate where Newton's equations had worked, his tensor-field equations reduce precisely to Newton's equation. Thus, without evoking the action-at-a-distance idea, Einstein's field formalism made the identical successful predictions in the way that Faraday had originally an-However, because Einticipated. stein's equations were more general than Newton's equation, it was implied that the tensor equations should be able to make additional predictions about gravity that are not at all implied by the Newtonian theory. This was indeed the case. The three predictions of his equations that were not previously made were: a frequency shift of monochromatic light as it propagates through a gravitational field; the bending of the trajectory of light as it passes through a gravitational field, and a precession of the axes of a planetary orbit. Each of these effects was observed to be in very good quantitative agreement with the predictions of Einstein's field equations.

# Towards a unified field theory

But Einstein did not consider his equations as another theory of gravity. Rather, according to what we said earlier, he considered them to be not more than one step towards the construction of a unified field theory in which electromagnetism, nuclear forces, weak interactions (and any other phenomena not yet discovered) might be incorporated with gravitation in terms of one geometrical field that expresses most generally the relation between points of the spacetime language to describe the laws of nature. He also anticipated that if such a unification could eventually be achieved, then perhaps the peculiarlooking consequences of the equations of quantum mechanics (for example the appearance of a wave-particle duality and nondeterminism) may emerge from an approximation for deterministic unified field equations.

Although both Faraday and Einstein anticipated a unified field theory to underlie the fundamental description of the material universe, Einstein extended Faraday's conceptual approach when he incorporated the principle of relativity. The implication in the studies of both Faraday and Einstein was that although it is convenient to utilize a space-time coördinate system to facilitate a mapping of force fields, space and time do not by them-

selves have objective connotation. However, a logical implication of the incorporation of the principle of relativity with Faraday's field concept is the appearance of physical significance in the relation between points of space-time as having to do with the mutual interaction of the matter content of the system. Thus a logical implication of the Einstein generalization of Faraday's approach is the necessary incorporation of the "test charge" with the field of force, to yield as the elementary entity one field of mutual interaction. A mathematical consequence of this generalization is that of passing from a formalism in terms of linear differential equations (such as Maxwell's equations) to a formalism in terms of nonlinear differential equations. Here we have a striking example in the history of science of passage towards increased conceptual simplicity (from one force field plus an arbitrarily defined test charge to a single field of mutual interaction) accompanied by decreased mathematical simplicity (from linear to nonlinear differential equations).

#### A GENERALIZED MACH PRINCIPLE

One of the important manifestations of interacting matter that I have not yet mentioned is inertia. Indeed, Einstein acknowledged that Ernst Mach's conclusions about the source of inertia had a major influence on his own theoretical studies of relativity theory. Let us then examine the conceptual notion that Mach introduced.

#### Mach's view of inertia

Mach's argumentation led to the conclusion that the inertial mass of any amount of matter (whether it be of microscopic, macroscopic or astronomical proportions) is not an intrinsic property of a thing. Rather it was argued that inertia is rooted in the dynamical coupling between the quantity of matter being studied and all of the other matter contained in the closed physical system that is described. Thus it is implied here that the most minute amount of matter, say an electron, has an inertial mass that is in fact a measure of its coupling to its entire environment, including the apparatus that measures the properties of the electron, the laboratory, the earth, the solar system, the galaxy, and so on, until all of the content of the universe is exhausted! Thus Mach rejected all previous views that assume that the mass of any bit of matter is one of its intrinsic properties.

When the principle that asserts this feature of inertia-the Mach principle -is incorporated with the theory of general relativity, it follows that inertial mass must be derivable from the field properties of space-time. Thus a full exploitation of the Mach principle implies that inertia must be incorporated with the other manifestations of interacting matter in the unified-field description. Einstein acknowledged the validity of this assertion but he did not reach the stage in his own studies where he would attempt to construct such a unification.

#### Derivation of inertial mass

How then can we set about discovering this relation between the field properties of space-time and the feature of inertia of matter? I have been taking the following approach in my own investigations of this problem.1 Let us start out by defining the inertial mass most primitively in terms of its appearance in the most fundamental equations that describe matter at the microscopic level. These were taken to be the field equations that have the Dirac form; they are firstorder differential equations in spinor variables, although they are not interpreted within this field theory as they are in the quantum theory, nor are they linear equations (see box on this page). Nevertheless, these matter field equations are constructed to approach asymptotically the linear eigenvalue form of the quantum-mechanical equations in the limit of sufficiently small energy-momentum transfer within the physical system. Normally one starts with this equation in special relativity and inserts a mass parameter. After the solutions of the equation have been found, the energy of the matter that is described can be computed. The resulting expression depends on the inserted mass parameter. One then adjusts the magnitude (and sign) of this parameter to fit the

On the other hand, my plan is not to introduce any mass parameter, but instead to express the matter-field equation in a curved space-time. When this is done, it turns out that with the most general type of spinor equation (in terms of the two-component spinor) a derived field appears in the place where the mass parameter was formerly inserted. The predictions of the theory then compel

one to identify this field with the inertial mass of the described matter. The derived mass field depends on the curvature of space-time. The latter geometrical property is, in turn, a manifestation of the mutual coupling of all of the matter within the closed system. Thus, if the rest of the universe should be depleted of all matter, the mass of the remaining electron, say, should correspondingly go to zero. The derived field relationship is then a quantitative expression of the Mach principle because here the inertial mass of any amount of matter

is indeed a well defined function of its dynamical coupling with all of the other matter within the entire closed system.

# Why is gravity only attractive?

To check this result further, it was found that the inertial mass field so derived is positive-definite. That is, with all possible changes in the metrical field, the particular combination of field variables that relate to the inertial mass is always positive. It then follows that even in the local limit, where the equations in special rela-

#### **SPINORS**

Dirac discovered that to express Schrödinger's wave mechanics in a relativistically invariant form, it was necessary to generalize the (complex) scalar field to a (complex) twocomponent field. This was called a "spinor variable" because of its relation to the previously discovered spin degrees of freedom that were empirically necessitated by the Zeeman spectrum of atoms. The spinor variable itself was actually discovered before Dirac's work by the mathematician E. Cartan. (Although the conventional expression of Dirac's equation is in terms of the four-component "bispinor" variable, these, in turn, are a union of two two-component spinors. The latter are the most primitive variables from which the Dirac formulation is built.)

After Dirac's discovery, Einstein and Mayer studied the following question: "Are the spin degrees of freedom a consequence of the postulates of the quantum theory or a consequence of the postulates of relativity theory?" To answer the question, they investigated the structure of the most primitive irreducible representations of the underlying group of relativity theory. They discovered that indeed the four-dimensional (real) representations (which describe the transformations of a fourvector between relatively moving observers) reduce to the direct product of two two-dimensional (complex) representations. The twocomponent (complex) functions that are the basis of these representations (analogous to the four-vector basis of the four-dimensional representations) were found to be the spinor variables that Dirac discovered to describe the electron!

Thus, Einstein and Mayer discovered the very important fact that the spinor variable is the most primitive expression of a relativistically invariant theory and that the spin degrees of freedom that appeared in Dirac's equation were actually a conse-

quence of relativity theory and not of the quantum theory. (Of course, once a relativistic theory is constructed in this, the most primitive form, it can then be "quantized" or not, depending on the particular theoretical study that is being pursued.) For this reason my study of inertia necessitated a formalism in terms of spinor variables in the most elementary description.

To mention one of the most important explicit features that define the spinor field variable, let us compare this two-component object with a two-component vector field in configuration space. If  $\bf A$  and  $\bf B$  are any two-dimensional (real) vector fields, then an invariant is the scalar product  $\bf A \cdot \bf B = A_1B_1 + A_2B_2$ . If

$$s = {s_1 \choose s_2}$$
 and  $t = {t_1 \choose t_2}$ 

are any two (complex) spinor field variables, the invariant combination is  $(s,t)=s_1t_2-s_2t_1$ . The components,  $s_1,s_2,t_1,t_2$ , are all continuous complex functions of the space-time coördinates. Note that this spinor invariant is then complex so that it actually corresponds to two invariants-the real and imaginary parts of (s,t). Finally it follows from the transformations between the spacetime coördinates of relatively moving reference frames in relativity theory that the spinor field variables also transform in a definite way from one frame to the others. Because of this result, and the fact that there is a maximal number of invariants in this theory (compared with all other relativistic formulations) this type of formulation predicts the maximal number of consequences. That is, a spinor theory will contain the same predictions as a vector theory, but it will make additional predictions that have no counterpart in the higher dimensional form. The electron spin coupling to an external magnetic field is an important example.

tivity have empirical validity, the inertial mass can have only one possible polarization. This result, in turn, implies that a gravitational force can have only one sign as this force depends on the product of the masses of the interacting matter. The latter result, of course, is in agreement with the experimental facts. To refute this result, it would be necessary to demonstrate the existence of negative mass, which is to say that we would have to observe the gravitational repulsion of one massive body by another, thereby implying the existence of masses with opposite polarities.

One other feature of the spinor matter-field equations that led to the preceding result is the necessary appearance of a coupling term corresponding exactly to the form of the electromagnetic coupling that eventually leads to the Lorentz force. That force, in turn, also depends within this theory on the metrical field, but in such a way that its form is nonpositive-definite; this term could be positive or negative, depending on the signs of certain derived functions of the metrical field. If one should define electromagnetic coupling most primitively in terms of its appearance in the matter-field equation that describes microscopic matter, then the conclusion must be drawn from this analysis that electromagnetic forces can be either attractive or repulsivea result that is also in agreement with the experimental facts. Thus we see that with the expression of the inertial features of matter in terms of the



DIRAC

primitive matter fields that solve relativistic spinor equations in a curved space-time (equations that asymptotically approach the equations of quantum mechanics) two important features of interacting matter, in all domains of interaction, are predicted from first principles to be in agreement with the experimental facts. These are the known properties that electromagnetic forces can be attractive or repulsive while gravitational forces are only attractive.

Finally, when one extends this analysis further it is found that the electromagnetic field intensity that corresponds to charged matter in motion is also dependent on the curvature of space-time and therefore on the matter content of the entire closed physical system.2 As in the case of inertial mass, the limit of a matterless universe corresponds here to an identical vanishing of the electromagnetic field intensity of any remaining bit of charged matter. This feature of the field equations of general relativity, which includes the inertial properties of matter, then relates to a generalized version of the Mach principle where all manifestations of interacting matter are in fact a feature of their dynamical coupling with the rest of the entire closed system.

#### **Implications**

The application of the generalized Mach principle necessarily implies that the universe is "closed", which means that there is no matter in the universe that, in an exact sense, is uncoupled (that is "free"). In principle, then, the same set of field equations that describe the domain of elementaryparticle physics must also describe the domain of laboratory dimensions and the domain of astronomical dimensions. The conclusion that follows is very difficult to believe at first glance; it is the possibility of discovering features about the structure of the universe in cosmological terms (the domain of light years) by studying the domain of elementary-particle interactions (the domain of fermis). For, according to the field theory that we have been discussing, there are indeed global features of the entire universe that appear in the quantitative description of elementary-particle interactions. One of these is the dependence of the inertial mass of an interacting elementary particle on the curvature of space-time. It is true, of course, that in such studies we are only

looking at an extremely small portion of space-time, corresponding to very high momentum transfer between matter and matter in the microscopic domain. Yet we are indeed looking at a portion of a continuous field that extends into the astronomical domain If we should have at our disposal the basic equations that determine these global fields, a knowledge of their behavior in the very small domain could well act as a boundary condition to extend our knowledge of the mapping of the metrical field to any domain of interaction. This ability would be analogous to predicting a large distortion on a radio wave front, being caused by a small bump on the emitting antenna, many miles away from the point of observation.

Another test of the implications of the generalized Mach principle would be to study the dependence of the electromagnetic field amplitudes on the curvature of space-time. The implication above was that as the effective curvature of space-time diminishes, the amplitudes of the electromagnetic force field correspondingly diminish. This effect can be tested by probing sufficiently small domains, which in turn correspond to sufficiently large quantities of energy-momentum transfer in the microscopic domain. The prediction here is that as the momentum transfer between matter in the microscopic domain approaches infinity, the electromagnetic coupling correspondingly approaches zero. It is as though the charge of interacting matter effectively reduces to zero at the origin. Experimentation on these effects is indeed within the domain of experimental high-energy physics studies that are presently in the planning stage in high-energy electron scattering programs. Current experimental studies on high energy electromagnetic scattering already hint at the validity of this theoretical conclusion.3

\* \* \*

This article is adapted from a lecture given to the physics club at the State University of New York College at Cortland

# References

- 1. M. Sachs, Nuovo Cimento 53B, 398
- 2. M. Sachs, Nuovo Cimento 55B, 199 (1968).
- K. W. Chen, A. Cone, J. Dunning, N. F. Ramsey, J. K. Walker, R. Wilson, in Nucleon Structure, eds. R. Hofstadter, L. I. Schiff, Stanford University Press (1964), p. 55.