Time in Relativity Theory: Arguments for a Philosophy of Being

OLIVIER COSTA DE BEAUREGARD

Translated and adapted from the French by David Park

The fundamental discovery of special relativity theory is that experimental facts admit a joint definition of time and length measures which entails a physical equivalence[1] between them. The present essay will review the historical development of this idea and its consequences for physics, and give a speculative discussion of the profound influence that it is likely to have on our philosophical views of the world.

1 *The Relativity Principle of Classical Dynamics*

The relativity principle of classical dynamics was a true predecessor of the special relativity principle of Einstein, and it was already related in many respects to fundamental aspects of the time problem.

The formulation of this pre-Einsteinian version of relativity lies between the absolute space principle postulated by Newton, and what may be called the relative motion principle of classical kinematics. According to Newton's absolute space principle[2] there must exist an absolute spatial reference frame relative to which all movements can be thought of as taking place. This idea turned out later to be metaphysical in character, *i.e.*, deprived of operational support. By stating this principle Newton gave a sort of formal status to common sense feeling; it may be that the postulate had its motivational root in the common experience of living on solid ground.

In complete contrast with the absolute space principle, the relative motion principle of classical kinematics seems at first sight to be experimentally established. This new principle emerged during the development of classical kinematics; it follows directly from Euclidean geometry and Newton's principle of an absolute time.[3] It states that *any* two solid reference frames, in whatever relative motion (translational acceleration, rotation, or arbitrary motion), are kinematically equivalent for the description of movements.

The relativity principle later discovered in classical dynamics is neither of these two, but lies, so to speak, between them. On the one hand, there is nothing in dynamics to substantiate the idea of an absolute reference frame but, on the other hand, dynamics gives a precise way to characterize absolute accelerations or rotations[4] (which classical kinematics cannot do). Thus, according to dynamics, the class of fundamental reference frames of space is neither as restricted as Newton supposed it to be, nor as broad as the purely kinematical relative motion principle would have it. One deduces in classical dynamics that the class under consideration is restricted to solid reference frames all in uniform relative translation with respect to each other; experimentation then allows the full characterization of these so-called Galilean frames.[5]

The simplest operational characterization of the class of Galilean frames refers to the inertial motion of a point particle. But, as Thomson and Tait have stressed in their famous *Natural Philosophy*,[6] this implies a simultaneous operational characterization of what may well be called a Galilean time scale t, for it is obvious that a point motion which is rectilinear and uniform, when referred to any Galilean frame and to a Galilean time scale t, will generally not remain so if referred to a non-Galilean frame and/or to a non-Galilean time scale $\tau = F(t)$. Therefore (and this is an important point for our purpose) *it turns out that there is a very close connection between the appropriate physical definition of a time scale and the Galilean relativity principle.*[7]

Natural clocks, that is, clocks evidently displaying Galilean time, may be any kind of inertial motion or, more generally, any motion implying the fundamental Newtonian formula of dynamics; such motions form the physical basis of mechanical clocks of either astronomical or laboratory size.

Now we must discuss the question of the measurability of time. My point of view will perhaps become clear if we briefly review the analogous development of a quantitative scale of temperatures. In the theory of heat it is found that temperatures are rendered *measurable* through the introduction of Kelvin's thermodynamic scale, or at least through the laws, due to Boyle and Gay-Lussac, for the compressibility and dilation of perfect gases. The point is that, before these definitions, the additivity of two temperature intervals could not be validly defined. But the discovery of the *universal laws* of perfect gases and/or of phenomenological thermodynamics created a new situation, permitting a universal relation of temperatures to other physical quantities; namely, to volume and pressure ($PV = nRT$) by use of a "perfect gas" thermometer, or, even better, to the mean value of the kinetic energy of (monatomic) gas molecules through the Maxwell-Boltzmann formula $\frac{1}{2}mv^2 = \frac{3}{2}kT$. The presence of the universal constant R or $k = R/N$ (N, Avogadro's Number) in the perfect gas formulas is significant, as it expresses a *physical equivalence* between temperatures and pressure-volume products or kinetic energies.

A very similar situation occurs in the time problem. Before the discovery, by Galileo and by Newton, of the universal laws of inertia and of inertial response to forces, the physical status of clocks was quite similar to that of thermometers before Lord Kelvin: there was no possible guarantee that a unique and valid definition of a time scale could be extracted from the performance of, say, sand or water clocks, or even from astronomical clocks.[8]

So, the new *universal law* (contrasted to the previous multiplicity of uncongruent physical clocks) may well be taken as Newton's formula for a point particle, $F = m\,\frac{d^2x}{dt^2}$. In this formula there is of course a universal constant present, but this constant is traditionally taken as equal to one with the dimension *zero* through our appropriate joint definition of the units of force, mass, space and time.[9] While the additivity of space intervals (in Euclidean geometry), of forces (through arguments based on statics) and masses (quantity of matter[10]) is taken as obvious, the additivity of Galilean time intervals as expressed in Newton's formula is established through the universal character of this formula.

An alternative statement is that the universal constant implicit in the Galileo-Newton formula establishes the physical equivalence between forces and mass-acceleration products—an equivalence which is directly experienced in the form of "inertial forces."

Our conclusion is that the Galileo-Newtonian universal laws of inertia have rendered time "measurable" in very much the same way that the universal laws of thermodynamics have rendered temperature measurable.

The profound significance of this remark is to be found in Einstein's and Minkowski's special theory of relativity.

2 *The Relativity Problem in Classical Optics and Electrodynamics*

When it had become clear that neither classical kinematics nor classical dynamics were able to define by themselves an absolute reference frame, it was hoped that studies in some other branch of physics would circumvent this apparent failure. In this regard kinematical optics, *i.e.*, the optics of moving systems, seemed at first quite promising. Indeed, the nineteenth-century physicists believed that the Huygens-Young-Fresnel optical waves were propagated in some appropriate medium which they named the "luminiferous ether"; and this hypothetical medium seemed likely to take the place of Newton's hypothetical absolute space. For example, according to classical kinematics, the spherical waves emitted at velocity c by a point source at rest in the ether would be expected to have velocities ranging between $c + v$ and $c - v$ in a reference frame moving with velocity v $(v < c)$ relative to the ether. So began the long history of the physical connection between kinematics and optics, the conclusion of which was to be Einstein's remodeling of kinematics after the requirements of electromagnetic theory.

In 1818, Arago proposed to detect the earth's "absolute motion" by measuring the refraction of starlight by a prism. This was a turning point in the history of physics, though it is clear today that Arago's way of questioning nature was not the most unambiguous one; Angström later improved the Arago test by using a source, a receiver, and a prism all at rest in the laboratory, so that no problem of a relative motion between the source and receiver was implied.

Nature's answer to Arago's question was negative: the observed refraction was the same as if the source, the receiver and the prism were all at rest relative to the ether. This came as an intellectual shock. Fresnel's answer to the riddle, known as the "ether drag postulate," was extremely remarkable: the formula was so adjusted that the effects of velocity v relative to the ether were eliminated up to the second order in $\beta = v/c$.[11] Thus the problem of finding a second-order effect was implicitly raised. When Veltmann[12] and Potier[13] had produced a theorem showing that due to Fresnel's formula the absence of first-order effects is absolutely general, the problem of finding a second-order effect was explicitly raised and this of course was the prologue to the famous Michelson-Morley experiment.

Before we come to this experiment, however, some more thinking on the state of affairs resulting from the work of Arago, Fresnel, Veltmann and Potier will yield profound insight into the relativity problem as viewed from its optical side. Our discussion will approach relativistic kinematics in a post facto way which is perhaps unfamiliar to some readers but which allows us to stress the most essential aspects of the subject.

It must be noted first that Fresnel's answer to Arago's result was metaphysical in that its wording still implied the notions of an ether and an "ether wind," while its formula was precisely built so as to eliminate (in the first order) all observable effects of the ether wind. A quite parallel situation arose later, but this time in the second order, with the "contraction hypothesis" that was Fitzgerald's and Lorentz's answer to Michelson's negative result.

Moreover, Potier made it clear that the Fresnel formula expresses a purely kinematical law of universal character,[14] namely, a composition law between three relative velocities: light vs. refracting medium, refracting medium vs. laboratory frame, and light vs. laboratory frame. This feature of the first-order Fresnel formula closely parallels that of the second-order Fitzgerald-Lorentz formula.

In 1908, von Laue showed that the Fresnel formula is merely a special case of the relativistic velocity composition law. The reciprocal step was taken in 1952, when Abelé and Malvaux[15] showed that if the Fresnel formula (in Potier's form) is postulated as the infinitesimal composition law of a group, the Einstein-Minkowski kinematics can be deduced.[16]

The group concept, of course, has been historically,[17] and is still essentially, one of the foundation stones of relativistic kinematics. But, in the seventies, the concept was hardly available to physicists; so the whole story had to be re-enacted in a strikingly parallel fashion, in the case of the second-order ether-wind effect.

In 1878, Michelson and Morley applied their interferometer to the problem of finding the supposed second-order effect of the ether wind. Once more no such effect appeared. Once more theoreticians formulated an ad hoc hypothesis: the Fitzgerald-Lorentz hypothesis, implying a universal formula of longitudinal contraction of material bodies under the ether wind. Once more there was something "metaphysical" in the discourse, the postulated ether wind and absolute frame of reference having no experimental counterparts. And once more the proposed formula was of a universal character, and purely kinematical in its nature.

In the meantime, two important concepts had come to maturity, whose association constitutes the key of the problem. On the one hand, various thinkers, among them Mascart[18] and Poincaré,[19] had become convinced that the relativity principle of dynamics is in fact the universal relativity principle, valid in all branches of physics. They concluded that the law of preferential equivalence of all Galilean frames is in fact a kinematical law which is also valid in optics, electrodynamics, etc. On the other hand, continuous group theory had emerged as a doctrine.[20] In essence, relativistic kinematics follows as a consequence of the application of group theory to optics or electrodynamics.[21]

It was the young Einstein[22] who showed that the unobservability of ether effects leads to a new joint definition of the length and time measures according to which the velocity of light is found to be the same in all reference frames. These measures are precisely those implied in the Lorentz transformation formulas connecting inertial frames[23]

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}},\ y' = y,\ z' = z,\ t' = \frac{t - vx/c^2}{\sqrt{1-\beta^2}}$$

($\beta \equiv v/c$) which read in reciprocal form

$$x = \frac{x' + vt'}{\sqrt{1-\beta^2}},\ y = y',\ z = z',\ t = \frac{t' + vx'/c^2}{\sqrt{1-\beta^2}}.$$

The main difference between the new Lorentz group and the corresponding classical Galileo group

$$x' = x - vt,\ y' = y,\ z' = z,\ t' = t,$$

$$x = x' + vt',\ y = y',\ z = z',\ t = t',$$

is that the transformation affects not only the spatial variables, but also the time variable t; thus, *a "proper" time t is attached to each inertial frame* or, in other words, *an inertial frame is not merely a spatial reference frame (as in the Galilean case) but also a temporal reference frame.* To emphasize this important difference, inertial frames are called Lorentzian rather than Galilean in the relativistic kinematics. It is well known, and obvious, that the limiting form of the Lorentz formulas when one lets $c \to \infty$ is the Galileo formulas; it may thus be said of the new kinematics what is written of the New Testament: that "it does not destroy, but it fulfills the Old."

By setting t or $t' = 0$, one verifies easily that the Fitzgerald-Lorentz contraction is built into the Lorentz formulas; the point is that, in the Einsteinian presentation, this contraction is reciprocal. Each of two

Lorentzian observers finds that the yardstick carried by the other seems shorter than his own.[24] This conclusion seemed highly paradoxical in its day; but there is in it no more mystery than in the well-known Euclidean fact of mutual foreshortening of distant objects known to be of equal size. A somewhat similar situation relates to time measurements, but here the situation is especially interesting because of an apparent paradox which has no analogue in the contraction of length. In picturesque form, the paradox considers a pair of twins, one of whom is an astronaut and undertakes a long space voyage while his brother, an administrator, stays home. Since the astronaut's time scale is contracted by the motion, it is expected that the astronaut should be biologically younger than his brother on his return. On the other hand, if there is no such thing as absolute motion, one might at first think that the astronaut could be taken as the reference point, with respect to which the administrator travels and returns. Then by the same argument as before, the administrator should be the younger when the journey is over. This "twin paradox" has been much discussed, but its resolution is basically simple. The doctrine of relative motion applies only to uniform motion in a straight line, and the astronaut is distinguished from his brother by the acceleration he undergoes. The astronaut is finally the younger of the two.

This closes our survey of the problem of kinematical optics. Before we start discussing Minkowski's remarkable interpretation of Einstein's theory, it will be useful to explain how, according to Duhem's and Poincaré's epistemological views,[25] Michelson's experiment "allows and suggests" the new relativistic joint definition of length and time measurements.

3 *Operational Commentary on the Results of the Michelson-Morley Experiment*

In this section we shall show that there exists a close connection between four classes of experiments: 1) experiments of the Michelson-Morley type; 2) optical measurements of length;[26] 3) Hertzian measurements of time[27] and 4) measurements of the speed of light. These considerations suggest that electromagnetic waves were truly predestined to furnish the scales for distance and time. They also permit us to understand without mathematics that to use these optical or Hertzian scales makes the speed of light in vacuo an invariant by definition.

1) Michelson's interferometer is essentially an optical scale arranged

so as to measure changes of length by detecting differences in the number of standing light waves along two rigid scales. It follows that the most direct interpretation of the negative result of Michelson and Morley is that the number of wave lengths of light emitted by a monochromatic source at rest with respect to a rigid scale and spread out along the scale is independent of the orientation of the scale. In 1887, when it was first firmly established, this result seemed highly paradoxical. But ever since the development of quantum mechanics in 1925–26, a positive result of the Michelson-Morley experiment would have seemed equally paradoxical. In fact, quantum mechanics describes any solid body as a standing de Broglie wave of complex structure,[28] and since waves of light are considered in quantum mechanics as a special case of matter waves, there is clearly no reason to suppose that the two kinds of waves should exhibit different kinematical behaviors.[29]

2) Michelson's negative result is required by the theory and practice of the optical measurement of lengths. If it were possible to detect the so-called "ether wind," any comparison between a rigid scale and an optical wave length would have to be preceded by a determination of the direction and velocity of the wind.[30]

3) Suppose on the other hand that one decided to use the period of a monochromatic optical radiation as a time scale. This raises no kinematic problem analogous to that just discussed. Now the problem is a dynamical one, for we must know whether the new time scale is (in the nonrelativistic limit) identical with that furnished by a body in uniform motion. Quantum mechanics again gives an affirmative answer: it is well known that the inertial motion of a particle is unambiguously described through a monochromatic plane wave[31] whose mathematical description is manifestly covariant in character.

We have thus explained how Michelson's negative result permits and suggests that we adopt at the same time the wave length and the period of a monochromatic optical wave as our standards of length and time.[32] Relativity thus legitimizes the situation that exists, and we can understand that it is the forms of the equations of d'Alembert and Klein-Gordon,[33] together with the Lorentz group under which they are invariant, that renders optical wave lengths and periods the natural measures of space and time. An optical wave is chosen rather than a matter wave because of the properties implied by the simpler equation of d'Alembert.

4) But to adopt the wave length and period of an optical wave as our standards of length and time is ipso facto to declare that c is an absolute

constant, a coefficient of equivalence between space and time. This is because of the exact relation

$$\lambda = c\tau$$

between the wave length λ and the period τ.

Finally, we may remark that a number of the most modern determinations of c follow the above conceptual scheme very closely. In the microwave cavity measurements of Essen[34] and of Hansen and Bol,[35] the spatial dimensions of the wave are determined by measurements of the cavity (some of them optical), while the period is compared with astronomical time. In the band-spectrum method of Plyler *et al*[36] and Rank *et al*,[37] c was determined by measuring separately the periods and wave lengths of the same molecular spectral lines.

At its first appearance, relativity seemed to mark a victory of optics and electromagnetism over mechanics. This was because neither kinematics nor dynamics had up to this time recognized the importance of the constant c. By now, not only the kinematics given us by relativity but also the dynamics given us by de Broglie, Heisenberg and their followers have assimilated into the physics of waves the constant c in an essential way. Today it would be possible to deduce all of relativistic kinematics, not from electrodynamics and optics via d'Alembert's equation, but more generally from the properties of matter described by the Klein-Gordon equation.[38] All is therefore once more in traditional order, with the quantum theory of the electromagnetic field a special case of a more general theory of mechanics which is the theory of quantized fields.

4 *Space-Time "Equivalence" and Minkowski's Four-dimensional Geometry*

The relativistic "equivalence" between space and time is strongly suggested by the elementary expression for a field component of a light wave

$$\phi = \cos 2\pi\left(\frac{t}{\tau} - \frac{x}{\lambda}\right)$$

where τ and λ are the wave's period and length, or by the form of d'Alembert's equation, of which ϕ is a solution. But the exact nature of the "equivalence" is expressed only by the Lorentz formulas which transform the spatial coordinates x, y, z, and the time coordinate t so as to leave invariant the quadratic form

$$s^2 = x^2 + y^2 + z^2 - c^2t^2.$$

It was Minkowski who established relativistic kinematics in a canonical form by showing that the Lorentz transformations can be regarded as rotations of a set of four Cartesian axes in a four-dimensional space-time with pseudo-Euclidean metric.[39] To verify this, we need only express the parameter β of the Lorentz formulas in terms of a new "angle" θ by

$$\beta = \tanh \theta$$

whence

$$\frac{\beta}{\sqrt{1-\beta^2}} = \sinh \theta, \qquad \frac{1}{\sqrt{1-\beta^2}} = \cosh \theta,$$

and the Lorentz formulas become

$$x' = x \cosh \theta - ct \sinh \theta, \qquad x = x' \cosh \theta + ct' \sinh \theta,$$

$$ct' = ct \cosh \theta - x \sinh \theta, \qquad ct = ct' \cosh \theta + x' \sinh \theta,$$

$$y' = y, \qquad z' = z.$$

These are indeed the formulas for a "hyperbolic rotation" in the plane whose coordinates are $x^1 = x$, $x^4 = ct$; this is a transformation which leaves invariant the "hyperbolic distance" $\sqrt{x^2 - (ct)^2}$ together with the coordinates $x^2 = y$, $x^3 = z$. The relation between the original and the transformed axes is shown in Figure 1.

One can go even further and force the geometry of space-time to be formally Euclidean by taking the fourth coordinate to be imaginary,

$$x^4 = ict, \; (x^4)^2 = -(ct)^2 \; (i = \sqrt{-1}).$$

The Lorentz formulas can now be written as

$$x' = x \cos i\theta + ict \sin i\theta, \qquad x = x' \cos i\theta - ict' \sin i\theta.$$

$$ict' = ict \cos i\theta - x \sin i\theta, \qquad ict = ict' \cos i\theta + x' \sin i\theta.$$

These are in the form of the ordinary expressions for the rigid rotation of a pair of axes, except that the coordinates are x and ict and the angle of rotation is expressed as $i\theta$.

Since the metric is pseudo-Euclidean, the cone defined by $s^2 = 0$ divides the directions of space-time into three classes (Figure 2): the exterior of the cone, $s^2 > 0$, and the two interior regions with $s^2 < 0$. Since these classes are distinguished by the value of the invariant s^2, no Lorentz transformation can take a vector from one class into another. The directions corresponding to the three classes are respectively called space-like ($s^2 > 0$), future time-like, and past time-like, both of the latter with $s^2 < 0$. Because the numbers x, y, z, and t are always real, it follows that the axes of x, y, and z are space-like while the t-axis is time-like. If we require the Lorentz transformations to be continuous,[40] all positive time axes will point into the future half cone, however they may

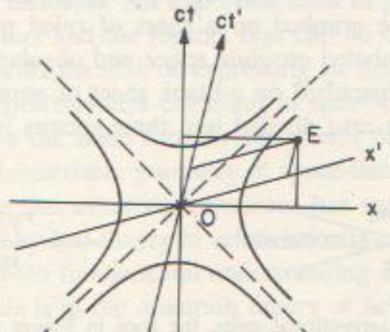

Figure 1. The axes x, representing a spatial direction, and ct, representing time measured in units compatible with those of x, can be used to represent the event E in a certain observer's space and time. For another observer moving relatively to the first, the coordinates of the same event are measured by the inclined axes x' and ct'. The hyperbolas are lines, or in general surfaces, which are described identically by both observers.

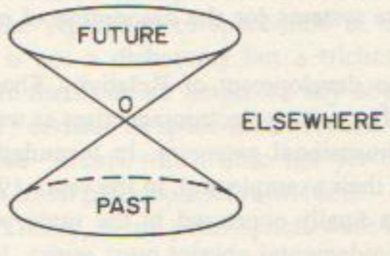

Figure 2. The double cone, shown with time dimension vertical and spatial dimensions horizontal to it, is generated by signals traveling inwards to and outwards from O with the speed of light. It divides space-time into three regions labeled *future*, *past* and *elsewhere*, which would be labeled identically by any observer, traveling with any uniform velocity, who coincided momentarily with O.

be transformed; the so-called "orthochronous Lorentz group" is thus defined.

We may note incidentally that expressed in ordinary units of space and time, the value of c is very large. The cone is thus very flat, and the *elsewhere* region is crushed between the *past* and *future* regions. Classical kinematics is obtained in the limit $c \to \infty$, which annihilates the *elsewhere* region (Figure 3). It is important to realize that the difference between classical and relativistic kinematics is not in the use of a four-dimensional space-time, but rather in the metric character first ascribed to this space by Minkowski.[41] To use a two-dimensional analogy, one may say that classical kinematics can be graphed on a sheet of ruled paper in which the coordinate axes are labeled *absolute space* and *absolute time*, while relativistic kinematics is inscribed on a blank sheet of paper, each point provided with a compass-card divided into three sectors labeled *past*, *future* and *elsewhere*.

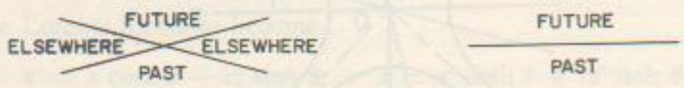

Figure 3. Measured in conventional units, the cone in Figure 2 is very flat and the *elsewhere* region is very narrow. In the Newtonian limit, in which the finiteness of the speed of light is ignored, this region disappears altogether and space-time consists only of the future and the past, separated by the present instant.

We have seen earlier that the essential new idea involved in the transition from pre-relativistic to relativistic mechanics is that the equivalence of Galilean (now Lorentzian) frames of reference is a law not only of dynamics but of kinematics as well. In Minkowski's geometry the interpretation of this point is very clear: it is the privileged equivalence of strictly Cartesian coordinate systems for the description of phenomena, just as in Euclidean space.

Very early in the development of Relativity Theory, Minkowski and von Laue succeeded in writing electromagnetism as well as dynamics in the language of four-dimensional geometry. In formulating wave mechanics, de Broglie followed their example and, in the years 1948–49, the quantum theory of fields was finally expressed in the same way.[42] It is a goal to which all parts of fundamental physics must aspire, for example, classical statistical mechanics.[43]

The theory of gravity poses a special problem. Though there have been, and are still, many attempts at a Minkowskian theory of gravity, the most successful theory is undoubtedly Einstein's General Theory of Relativity

(1915), which is expressed in terms of a Riemannian geometry in four dimensions. Here the geometry of Minkowski is only a locally tangent approximation, as a tangent plane may locally approximate a curved surface in the geometry of ordinary experience.

In any theory whose mode of expression is geometric—and we have seen that relativity theories are of this kind—a very important concept is that of covariance. What is objective in a geometric theory is what is defined independently of the way in which coordinates are assigned to the space, for example, points, lines, surfaces, etc., together with the figures formed from these elements. An important class of geometric "objects" is formed by the vectors and the tensors that can be defined at each point. We owe to Minkowski the idea of expressing all the fundamental laws of physics in the form of relations covariant in space-time, that is, tensorial relations which take the same form in all systems of coordinates. Physics is thus related to the intrinsic geometry of space-time. Looking back over history, we can see that every time someone has succeeded in expressing the problems of a physical theory in relativistically covariant terms, it has led to advances in both fundamental understanding and technical skill; the most recent example is in the quantum theory of fields.[44] For this reason, the concepts of physical objectivity and of covariance have become nearly synonymous in Relativity Theory; that is, in a strictly relativistic discussion only those objects and relations which can be expressed covariantly can be considered as objective properties of the world.

As concerns philosophy and general culture, perhaps the most important consequences of the requirement of covariance concern the relation between past and future. In Newtonian kinematics the separation between past and future was objective, in the sense that it was determined by a single instant of universal time, the present. This is no longer true in relativistic kinematics: the separation of space-time at each point of space and instant of time is not a dichotomy but a trichotomy (past, future, elsewhere). Therefore there can no longer be any objective and essential (that is, not arbitrary) division of space-time between "events which have already occurred" and "events which have not yet occurred." There is inherent in this fact a small philosophical revolution.

Before Relativity Theory, many philosophers were inclined to consider matter as occupying a certain region of space but, in respect to time, as being concentrated in an instant without extension; this view was compatible with the kinematics of Galileo and Newton, but is incompatible with that of Relativity Theory. If matter has spatial extension, it follows (in virtue of the trichotomy mentioned above) that it has also extension in time.[45] This is why first Minkowski,[46] then Einstein,[47] Weyl,[48] Fantappiè,[49] Feynman,[50] and many others have imagined space-time and its material contents as spread out in four dimensions. For those authors, of whom I am one, who take seriously the requirement of covariance, relativity is a theory in which everything is "written" and where change is only relative to the perceptual mode of living beings. Humans and other living creatures, for reasons which one can try to explain,[51] are compelled to explore little by little the content of the fourth dimension, as each one traverses, without stopping or turning back, a time-like trajectory in space-time.

We can, of course, easily imagine the past experience of an individual as being diagrammed in space-time along a time-axis extending backwards into the past. In our view, the future is to be adjoined to the diagram in a corresponding fashion. There are writers who affirm that the future contains elements which are undetermined by the past and the present, and that the future light cone in the diagram must therefore be left blank. The answer is that there *is* a future. Nature "will take" one of the alternatives open to her, and it is this that we must imagine inscribed, even though we do not know what it "will be."

5 *Causality and Free Will: A Relativistic Formulation of the Problem*

When Magellan's companions reappeared from the East after having set out towards the West, a group of interwoven problems concerning the sphericity of the earth and the heliocentric theory of Aristarchus and Copernicus ceased to be academic; they now concerned everyone. Currently, thousands of technicians in many branches of physics and engineering use the findings of relativity in their work. Perhaps the day is not distant when school children must learn to think in terms of Minkowski diagrams. This will be no harder for them than it is to accept that the earth is a spinning ball, small in the general scheme of things, slowly revolving around a gigantic sun. The only difficulty is to learn that the everyday phenomena of astronomy, even though they do not at once suggest to our minds a heliocentric system, can nevertheless easily be explained in this way. We learn to do this as children under the guarantee that the heliocentric system is in harmony with the general principles of nature (*i.e.*, the laws of mechanics) while the geocentric system is not.

We encounter the same kind of difficulty in trying to adjust our intuitive views of time and space to the relativistic view, but the same guarantee encourages us to make the attempt.

Perhaps the most obvious points at which some effort must be made to harmonize relativity and intuition are the apparently one-way flow of time and the related problem of free will. We shall briefly discuss both.

The most conspicuous signs of the unidirectionality of time can be traced to our participation in the general evolution of the universe. The universe (at least the region of it known to us) consists of a markedly unstable arrangement of matter and energy evolving irreversibly in the direction of equilibrium. This subject is discussed elsewhere in this volume, and we may say that it is reasonably well understood. This has interesting consequences in the domain of elementary processes, such as the emission and absorption of radiation, governed by equations which are relativistically invariant and yet in which the law of cause and effect seems to demand that a distinction be made between the past and future directions of time.

The nature of the question is illustrated by a simple example. Imagine a quiet pond with leaves in the water around the edge. Into the pond at the point P a stick is dipped and withdrawn. Waves travel outward from P towards the edge of the pond and when they arrive, move the leaves in the water. This is an example of cause and effect, chosen to put in evidence the wave mechanism by which the dipping of the stick causes the leaves to move. It is obvious in it that the effects follow the cause in time. Now, formally speaking, the entire process is reversible. If it were reversed, the leaves would move, a circular wave would detach itself from the bank and converge towards P, at the right moment the stick would dip in and out and the pond would then be silent. In this process the roles of cause and effect are reversed, and it is possible to show that the entropy of the universe would *decrease* as a result of it.

In this example we see the interconnection of three tendencies in nature: that waves move outward from a disturbance and not inward towards it, that effects follow their causes and do not precede them, and that natural processes increase entropy and do not decrease it. In the first and second of these, nature seems not to use the symmetry given it by relativistic laws: the third shows why this is so; even elementary processes, or at least almost all that we would wish to discuss, are involved, in a probabilistic fashion, with the general increase of entropy. And since the process occurring in the pond is analogous to such processes as the radiation of energy by atoms and the interaction of charged particles via the electromagnetic field, we see that the irreversible evolution of the universe as a whole imposes, *via* probability, its dissymmetry even on the microscopic scale.[52]

We come now to the question of determinism versus free will. It is an exceedingly complex one, especially because it cannot be analyzed without a detailed use of quantum mechanics.

The subject concerns relativity because if a conscious being is able to make free decisions, these decisions influence the future and not the past and once again a dissymmetry exists. Further, it seems to be urgently implied that the future half of the Minkowski diagram cannot be filled in, even in principle, because decisions upon which it depends have not yet been made. To this we have answered above that *something* is going to happen and that is what is written down. This does not answer the question of determinism, however, and we must proceed to make a basic distinction.

There are at least two different ways in which we can imagine events to be inscribed on a Minkowski diagram: as a mathematical description in terms of the psi-function of the entire universe, and as a simple record of the sense impressions of one observer. We have a free choice. Theoretical physics imagines a universe which evolves subject to precisely causal laws except for processes in which conscious beings intervene. Such interventions are governed by essential indeterminacies whose result is that our successive sense impressions cannot in general be put in exact causal relation with each other. In this sense, causality is often said to have disappeared from physics and, in this sense, a four-dimensional diagram of an observer's sensory experience will not consist of causally related events. One of these two modes of description is causal, but it is not causally related to sensory experience. The other is a direct transcription of sensory experience, but it is not causal.

It should not be thought that the two possibilities represent different schools of thought or that they conflict with each other. Instead, they are in Bohr's[53] sense complementary ways of describing the same thing.

> Just as the freedom of the will is an experiential category of our psychic life, causality may be considered as a mode of perception by which we reduce our sense impressions to order. At the same time, however, we are concerned in both cases with idealizations whose natural limitations are open to investigations and which depend on one another in the sense that the feeling of volition and the demand for causality are equally indispensable elements in the relation between subject and object which forms the core of the problem of knowledge.[54]

Each view, the deterministic one and the one that emphasizes freedom of choice, has its appropriate mode of expression. In the *Heisenberg picture* of quantum field theory the psi-function, assumed to contain the completest possible specification of the universe, is independent of time. The equations governing the various interacting fields are formally the same as they would be in classical physics. The psi-function in any representation denotes which possible states are occupied and which are not. In the

Heisenberg picture the occupation numbers of the states never change: nothing happens. This picture is a reformulation in quantum terms of the classical and deterministic conception of nature in which everything is written once for all.

In the *interaction picture*, the various fields are described by equations which omit all interactions. The psi-function, on the other hand, reflects the effects of the interactions and evolves continuously in time. It represents the changing face of Nature as we know it, while the Heisenberg picture may be said to represent it in God's view. The crucially important point is that the two pictures are complementary descriptions of the same thing. A mathematical transformation enables us to pass from one to the other at will and to know that the physical content of the two pictures is exactly the same. This is an interesting example of how two viewpoints, which from a non-mathematical point of view may appear irreconcilable, may in fact be merely complementary. In one sense they are mutually exclusive, but both are necessary to our understanding of the Universe and its relation to man.

These, of course, are the author's own speculations on a difficult subject which has not yet been settled in a way generally accepted by all thinkers. They are given as an example of the perplexing problems which are raised as soon as the mind-and-body relations are thought of not "in space at a given instant," which would not be relativistically covariant, but in space-time.[55]

6 *Conclusions*

We may say that Special Relativity Theory has shed much light upon old riddles in kinematical optics and in innumerable other problems of contemporary physics. In this respect, Minkowski's theory of space-time has provided a powerful tool of far-ranging validity; perhaps future generations will say that Minkowski was a second Euclid, for he has made a theory of geometry marvelously describing the physical world as we know it today.

Relativity theory has drastically changed the concepts we use to describe the physical world, that is, in a broad sense, our cosmological view of things. This is a new Copernican or Magellanic revolution, the consequences of which are almost unexplored. Of these consequences we have tentatively submitted an example in the last section—a few other ones being also very interesting to consider even if they cannot be settled *hic et nunc*.