12 May 1972, Volume 176, Number 4035

SCIENCE

Time Reversal

The notion that the direction of time flow is not knowable appears to be upset by recent experiments.

Robert G. Sachs

The notion that the direction of flow of time cannot be determined by any physical experiment has been deeply ingrained in the thought and theories of physicists until very recent years. In all cases in which they have been adequately tested, the laws of physics satisfy this condition and, in unexplored areas, the condition is usually taken as a starting principle for formulating a theory.

These were the opening sentences of an article that appeared in *Science* in 1963 under the title "Can the direction of flow of time be determined?" (*1*). That article was written because new ideas and results of particle physics had suggested entirely new experimental and theoretical approaches to this venerable question. My motivation in taking up the question again is that it now appears to have been answered —and answered, unexpectedly, in the affirmative—by the results of an extensive series of experiments beginning with the landmark experiment of Christenson, Cronin, Fitch, and Turlay (*2*), who showed that the long-lived neutral K-meson could decay into two π-mesons. Evidently a sense of flow of time is determined by this series of experiments.

The implications for physical theories have yet to be understood, but the results of these experiments have a deep significance for physics and may be important to cosmology. It is my purpose here not to speculate on the theories but rather to try to put the experiments into perspective with regard to the general question of time reversibility. The connection between the behavior of elementary particles and time reversal is not a trivial one in general, but some of its aspects can be understood by direct examination of the motions of the particles. When dealing with such aspects, the physicist often refers to motion reversal rather than time reversal. However, it turns out that the significant phenomena that have been observed thus far are not manifested directly by particle motion but give much more indirect information concerning time reversal. It is part of my purpose here to show how this very indirect information may be made more intuitively understandable.

Our intuitive notions concerning the direction of flow of time are, of course, thoroughly entangled with our chemical, biological, and psychological experience with the irreversibility of time. These prejudices must be understood and isolated in order that the real question of time reversibility of the laws of physics may be understood. Some simple examples of time-reversal invariance in classical mechanics and their implications with regard to everyday experience show that there is no inconsistency between the reversibility of the laws of physics and the irreversibility of experience. They serve as a background for the discussion of the evidence for the irreversibility of certain laws of physics.

The impetus for examining the role of time reversal in particle physics was a direct consequence of the discovery (*3*) that the left-right symmetry of the laws of physics is violated by the weak interactions, that is, those interactions that are responsible for nuclear β-decay and the decay of elementary particles. This violation of the space-reversal symmetry of the physical laws suggested quite naturally an examination of the time-reversal symmetry. In fact, there are close connections between space reversal and time reversal from the point of view of the structure of physical theories. These connections are interrelated with a third operation, called charge conjugation (C), which replaces every particle of a system by its antiparticle, as will be discussed in more detail later.

Space reversal is another convenient starting point for my principal topic because, with the help of mirrors, it is relatively easy to visualize the dynamic effects of space reversal while it is not so easy to visualize consequences of time reversal.

Time Scales and Classical Motion

In the description of physical phenomena, the time variable enters at the most elementary level as a parameter serving to identify the order of a sequence of events. To give the variable quantitative meaning, a calendar (*4*) is constructed in terms of a well-established and continuing sequence of repetitive events. If the period of repetition is manifestly constant, it may be used as a unit of time. In order to complete the determination of a time scale, it is then only necessary to decide at what point to begin counting.

Such a scale is indicated in Fig. 1a. The labels B.C. and A.D. are needed in order to describe events in both the past and the future. However, this labeling system is awkward if the time scale

The author is professor of physics and director of the Enrico Fermi Institute of the University of Chicago. This article is based on his vice-presidential address at the 137th meeting of the American Association for the Advancement of Science in Chicago, Illinois, December 1970.

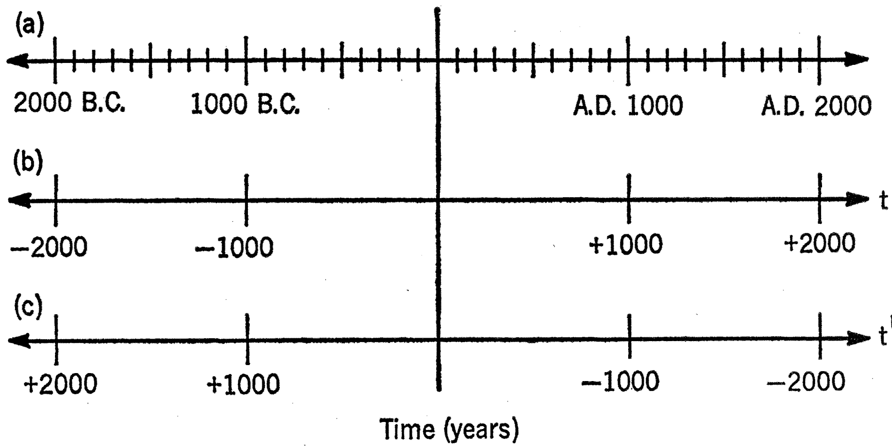


Fig. 1. Time scales established on the basis of (a) the conventional calendar, (b) a "normal" convention for defining positive and negative values of the time variable, and (c) the reversed convention.

is to be used in mathematical equations, equations that will be needed to describe the laws of motion of physical systems. Since the concept of time sequence is susceptible to addition and subtraction, the labels B.C. and A.D. should be replaced by opposite algebraic signs. For example, the algebraic time variable t may be defined as in Fig. 1b.

The point that is important here is that the choice associating B.C. with negative time and A.D. with positive time is arbitrary; it is merely a matter of convention. The convention is "reasonable," of course, because it agrees with our habits, but it is a convention. The time variable t', described in Fig. 1c, in which the opposite association is made, would also be acceptable for describing the motion.

The algebraic connection between these two equally acceptable time variables is

$$t' = -t \qquad (1)$$

When Eq. 1 is used to replace t by t', in mathematical formulas depending on t, the formulas are said to be subject to a time-reversal transformation, denoted by T. The formulas of interest here are the mathematical expressions of the laws of motion of physical systems, usually called the equations of motion. The question of interest here is whether these equations of motion are unchanged (invariant) when subjected to the time-reversal transformation. An equivalent question is whether the equations of motion of any physical system take a different mathematical form when the convention of Fig. 1b is replaced by the convention of Fig. 1c. When the question is phrased in this way, our intuitive reaction is that the physical laws should not depend on a mere convention; therefore, that they should turn out to be invariant under time reversal.

To help illustrate what this means, I shall make use of the concept of space reversal, which is more commonly called space inversion in physics. Space reversal concerns the choice between the right-handed and left-handed coordinate systems shown in Fig. 2, a choice that is again a matter of convention. The transformation of formulas between the two, made by use of the relations

$$\begin{aligned} x' &= -x \\ y' &= -y \\ z' &= -z \end{aligned} \qquad (2)$$

is called inversion and is usually denoted by P (for parity). The question here is whether the equations of motion are unchanged when subjected to the transformation P or, again, whether the choice of convention makes a difference in the laws of physics.

Since the viewing of any system in a flat mirror leads to an exchange of right- and left-handedness, the question may be rephrased in terms of a comparison between the laws of motion as observed in the laboratory and those that would be observed by making the measurements on the image of the system viewed in a mirror.

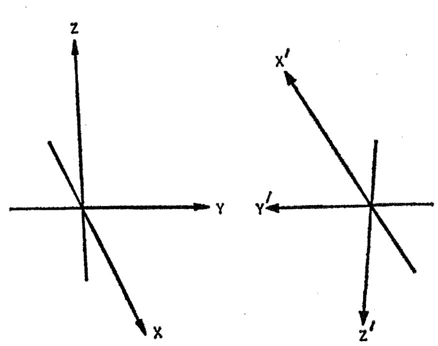


Fig. 2. Right-handed and left-handed coordinate system.

A simple example will serve to illustrate the point. Consider the motion of a particle moving under the influence of a massive center, shown as the heavy black dot in Fig. 3. If the motion in the laboratory is given by the orbit labeled *original* in the figure, then the space-reversed motion is given by the image of the orbit. The motion is not the same in the left- and right-handed systems, but that is not the question under consideration. The question is whether the laws of motion are the same. The important point is that the motion is not only determined by the laws of motion, but it also depends on the starting conditions (initial conditions), which are imposed externally and, therefore, are not subject to any general symmetry principle. A particle may be thrown from left to right or from right to left with equal ease. It will then be subject to the same laws of motion but will not follow the same motion.

Since the initial conditions (in this case, the initial position and velocity of the particle) are different in the image from those in the original, it must be expected that the motion is different even if the laws of motion are the same. For a comparison of the laws of motion rather than the motion itself, the experiment may be repeated in the laboratory with the same initial conditions that one sees in the mirror. The result is then the orbit labeled *actual reversal* in Fig. 3. This reversed orbit and the image orbit should be identical at every point if the laws of motion are invariant under space reversal.

The application of time reversal in place of space reversal can be illustrated by replacing the mirror used for space reversal by a motion picture run backward. The motion of a particle under the influence of a massive center is again used for demonstration in Fig. 4. The original orbit in the laboratory is shown in the top center of the figure. The dashed orbit is the result obtained when a motion picture of the original motion is run backward. Again, as in the case of mirroring, the motion is different in the two cases, and the reason is the same, namely, the initial conditions are different.

When the experiment is repeated in the laboratory with initial conditions matching those shown in the reversed movie, then the orbit is the one labeled *actual reversal* in Fig. 4. If the laws of motion are invariant under time reversal, then a point-by-point comparison of the orbits in the reversed movie and the reversed motion should show that they are identical.

These manipulations with motion pictures and mirrors may seem quite trivial and lacking in content, but the following example will serve to show how they could reveal new information about physics. Assume that you are familiar with permanent magnets but know nothing about the origin of magnetism. You construct a permanent magnet giving a uniform magnetic field perpendicular to the page, as shown in Fig. 5. If an electrically charged particle is shot from the side into this field, it describes the orbit shown in the top half of the figure.

You now carry out the time-reversing program described before by making a motion picture of the orbit and running it backward. The result is shown by the dashed curve in Fig. 5. You observe that in the reversed movie the initial condition calls for shooting an identical charged particle into the field from above. Therefore, in the laboratory, you shoot a particle in as prescribed and the orbit you will observe is not at all the same as in the movie; it is the solid curve labeled *motion reversal* in Fig. 5.

From this you might conclude that the laws governing the motion of a charged particle in a magnetic field violate the principle of time-reversal invariance. On the other hand, because of a desire to cling to symmetry with regard to the sense of time flow, you could surmise that the magnetic field of the permanent magnet must be produced by motion of particles within the iron rather than by stationary particles of a particular magnetic pole strength. The contradiction illustrated by Fig. 5 would then be ascribed to your failure to make the appropriate changes in the initial conditions on the motions of these particles.

Of course, we know that the second explanation is confirmed by other evidence; magnetic fields are generated by the motion of electric charges (in this case, the motion of electrons) and are reversed when the motions of the charges are reversed. There is no contradiction of T-invariance when the initial conditions on these currents are also taken into account in our description of motion reversal.

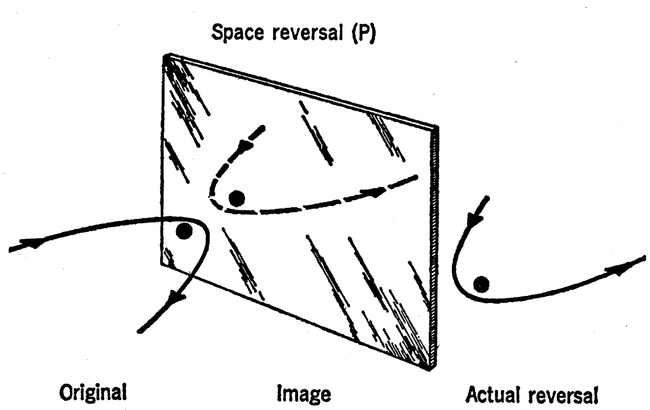
Complex Systems

I have emphasized the distinction between the role of the initial conditions and that of the laws of motion for simple systems. For complex systems, the role of the initial conditions becomes critical for understanding time reversal. In fact, the initial conditions are the determining factor in our understanding of the irreversibility of the chemical and biological phenomena that give us our natural sense of the direction of time flow.

The critical role of initial conditions can be shown even for a relatively simple system such as the four blocks shown in the motion picture sequence (Fig. 6a). The original motion is observed by reading the film strip in the forward direction: The stack of blocks is tipped and falls, and each block is caught by one of the actors. To test time reversal, we proceed as before. The motion shown by the reversed movie is indicated in Fig. 6b. To test time reversal, then, we must ask the actors to try to repeat what we see in the reversed movie, that is, to establish the initial conditions for reversed motion by throwing each block in precisely the way it appears to be thrown in Fig. 6b. We know that it would be very difficult for the actors to meet these precise conditions and throw the blocks so that they stack up, as in the reversed movie.

In fact, if we saw a motion picture showing the actors performing that stunt, we would immediately know that it was being run backward. As shown, the stunt simply would not be credible. The origin of this incredibility is our knowledge, gained from experience, that since there are such a vast number of ways in which the actors can throw four bricks, the probability of their finding just that one precise way needed for the stunt is extremely small. Thus, the incredibility may be equated to improbability.

For more complex systems, say molecular systems, there is much greater incredibility attached to the probability for reversal of a motion from an ordered to a disordered arrangement. In fact, we no longer perceive enough detailed information to know how the initial conditions of each molecule should be fixed in order to reverse the motion. We describe the system in terms of averages, such as the thermodynamic variables, over the detailed molecular motions, and the various possible motions are weighted by their probabilities in the averaging process. It is in this averaging process that the irreversibility of the motion is intro-



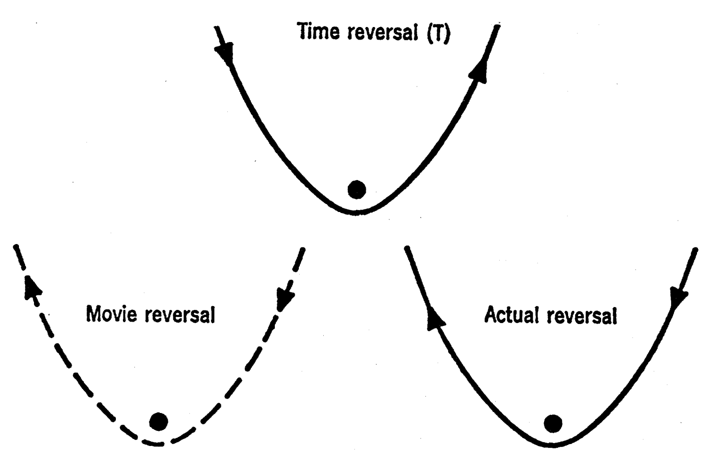


Fig. 3 (left). Orbit of a particle moving under the influence of a massive center as viewed in a mirror and as obtained after space reversal. Fig. 4 (right). Orbit of a particle moving under the influence of a massive center as viewed in a reversed movie and as obtained after motion reversal.

duced because the initial conditions on each molecule required for exact reversal of the motion have an incredibly small probability and, therefore, an incredibly small weight in the averaging.

This can be demonstrated by Gibbs' example in his discussion of coarse-grained as opposed to fine-grained averaging (*5*). He considered an incompressible fluid made up of two components in equal parts, one of them being opaque and the other transparent. Initially the two fluids are separated; let us say that the lower half of the container is occupied by the opaque fluid and the upper half by the transparent fluid. If the fluid is thoroughly stirred, it will appear to be neither opaque nor transparent but translucent. This is the coarse-grained view. If the microscopic or fine-grained structure of the stirred liquid is examined, it will be found to consist of threads of opaque and transparent liquids, the total volumes of the two being equal. Thus, the microscopic information is there, and exact reversal of the stirring would lead to unmixing of the fluid into its two components (*6*).

However, from our point of view, the reversibility of the motion is to be established by preparing the system with initial conditions matching those seen in a reversed movie. If only the macroscopic conditions needed to be matched, there would be no difficulty. There are an enormous number of ways in which the threads of liquid could be arranged to give the appearance of a uniformly translucent liquid. However, for the exact reverse motion to occur, it is necessary that a particular one of these many configurations be arranged. The a priori probability of arranging the threads in this way is incredibly small; hence, the mixing appears to be irreversible, although the laws governing the motion of the system may be perfectly reversible.

This information is usually expressed by the statement that certain gross properties of the system, such as its entropy, can change only in one way. Gibbs cited the example of the average of the square of the transparency of the liquid. If the clear liquid has transparency equal to 1 and the opaque liquid has transparency equal to 0, the average square in the unmixed state is $(1^2 + 0^2)/2 = 1/2$. In the mixed state the transparency in a coarse-grained observation is ½ throughout the liquid so that the average of its square is ¼. Of course, if the average of fine-grained observations could be made, it would be no different from the original case. One can show that for any kind of mixing (uniform or otherwise) the coarse-grained average is always less than ½. Therefore, one concludes that this quantity can only decrease with time, no matter what convention is used for the direction of time flow. Thus, from the coarse-grained, or thermodynamic, point of view the motion is irreversible, although the detailed motion of the threads is governed by reversible laws.

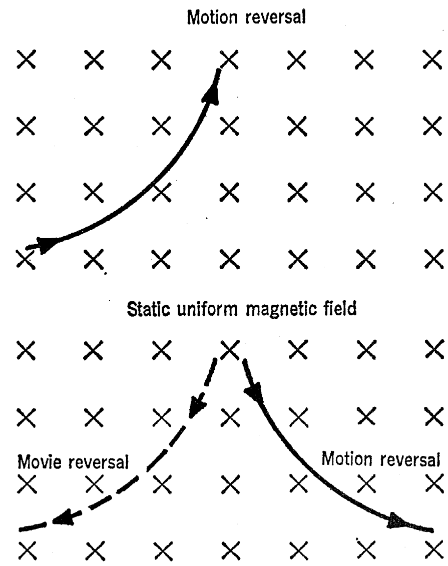


Fig. 5. Orbit of a charged particle moving in a uniform (permanent) magnetic field as viewed in a reversed movie and as obtained after reversal of the motion of the particle.

If the most complex of systems, the universe, is considered, it is easy to imagine that it started from an ordered condition; for example, a collection of protons distributed uniformly throughout space or possibly distributed uniformly on the surface of a sphere. The evolution of a simple model of the universe based on the latter assumption is illustrated schematically in Fig. 7. In Fig. 7a the evolutionary process is shown on the normal time scale t from gravitational collapse, the generation of elements through nuclear reactions, and the formation of chemical compounds to the creation of living organisms and their evolution. Figure 7b shows what would be seen in a motion picture of the evolutionary process run backward.

To duplicate the latter motion to test time-reversal invariance would require establishing in the laboratory the initial conditions described by the *etc.* in the figure on the t' scale. Because of the myriad of variables in the system, it is not conceivable that these initial conditions could be or could have been established. Therefore, the evolution of the universe, biological evolution, aging, and so forth, could only conceivably go one way, from the ordered to the disordered state. That gives us our sense of absoluteness of the direction of flow of time.

Weak Interaction, Parity Violation, and CP-Invariance

As we have seen, macroscopic systems are not suitable for tests of time reversal because the initial conditions play such an overwhelming role in determining the motions of the system. However, it is possible to test the principle by means of sufficiently simple systems of particles. Thus, tests have been carried out by studying the properties of specially selected nuclear systems and systems of elementary particles.

The validity of the principle of invariance under space reversal has also been tested with such simple systems. The only phenomena for which the failure of either symmetry principle has been established are those produced by the weak interactions between particles.

To explain what is meant by the term weak interaction, I make use of the notion that the laws of motion in physics are usually expressed in terms of the energies of interaction between parts of a system. When a system is reduced to its most elementary constituents, it is found that for all known systems there are only four kinds of interactions that need be considered. Two of these are the classical interactions of gravity and electromagnetism. For example, the electromagnetic interaction is responsible for the electric and magnetic forces acting between charges and electric currents, and it is also responsible for the emission of light by transitions between different energy states of atoms and molecules, since light is an electromagnetic phenomenon.

The other two fundamental interactions manifest themselves only in nuclear and elementary-particle phenomena. One is called the strong interaction, and it is responsible for the very strong forces that hold atomic nuclei together. The strong interaction also accounts for the production of elementary particles in high-energy collisions between nuclear particles and for a multitude of other properties of the particles of physics called hadrons.

The weak interaction is entirely different in character. Such small interaction energies are involved that it does not lead to any measurable forces, nor are these forces capable of holding anything together. However, like the electromagnetic interactions, they are capable of causing transitions between states of nuclei or elementary particles, with the emission of other particles. The form of radioactivity known as β-decay, in which a transition from one atomic nucleus to another takes place with the emission of an electron, was the first known reason to introduce such an interaction into physics. It turns out that the β-decay of all nuclei can be understood in terms of the weak interactions of their basic ingredients, the proton and the neutron, and that every other elementary particle (with the exception of the photon) is subject to weak interactions causing those in higher energy states to decay into others in lower energy states.

I have mentioned before that the impetus for examining the role of time reversal in weak interactions was produced by the discovery (*3*) that they violate space-reversal symmetry. The actual discovery was that, under the right conditions, the number of β-particles (electrons) emitted by a radioactive ^{60}Co source placed in a magnetic field is larger in the direction opposite to the field than in the direction along the field. That this result implies a violation of the left-right symmetry is made apparent in Fig. 8, which shows schematically the original result in the laboratory and its mirror image. In this figure, the wire loop is perpendicular to the mirror, producing a magnetic field B parallel to the mirror. The distribution in intensity of the β-particles in the plane perpendicular to the mirror is illustrated by the lengths of the arrows.

The procedure for testing space reversal, as set forth in connection with Fig. 3, is to repeat the experiment in the laboratory under the conditions seen in the mirror. The electric current in the loop has the opposite direction in the image. When this condition is reproduced in the laboratory, it will lead to a magnetic field in the opposite direction (physicists would say that the magnetic field vector has even parity, which is "unnatural" parity for a vector). But if the experiment is repeated in the laboratory under the reversed-field condition, the direction of maximum intensity of β-particles is also reversed. That can be seen by recognizing that the desired new configuration in the laboratory can be obtained by merely rotating the original apparatus by 180° about a vertical axis.

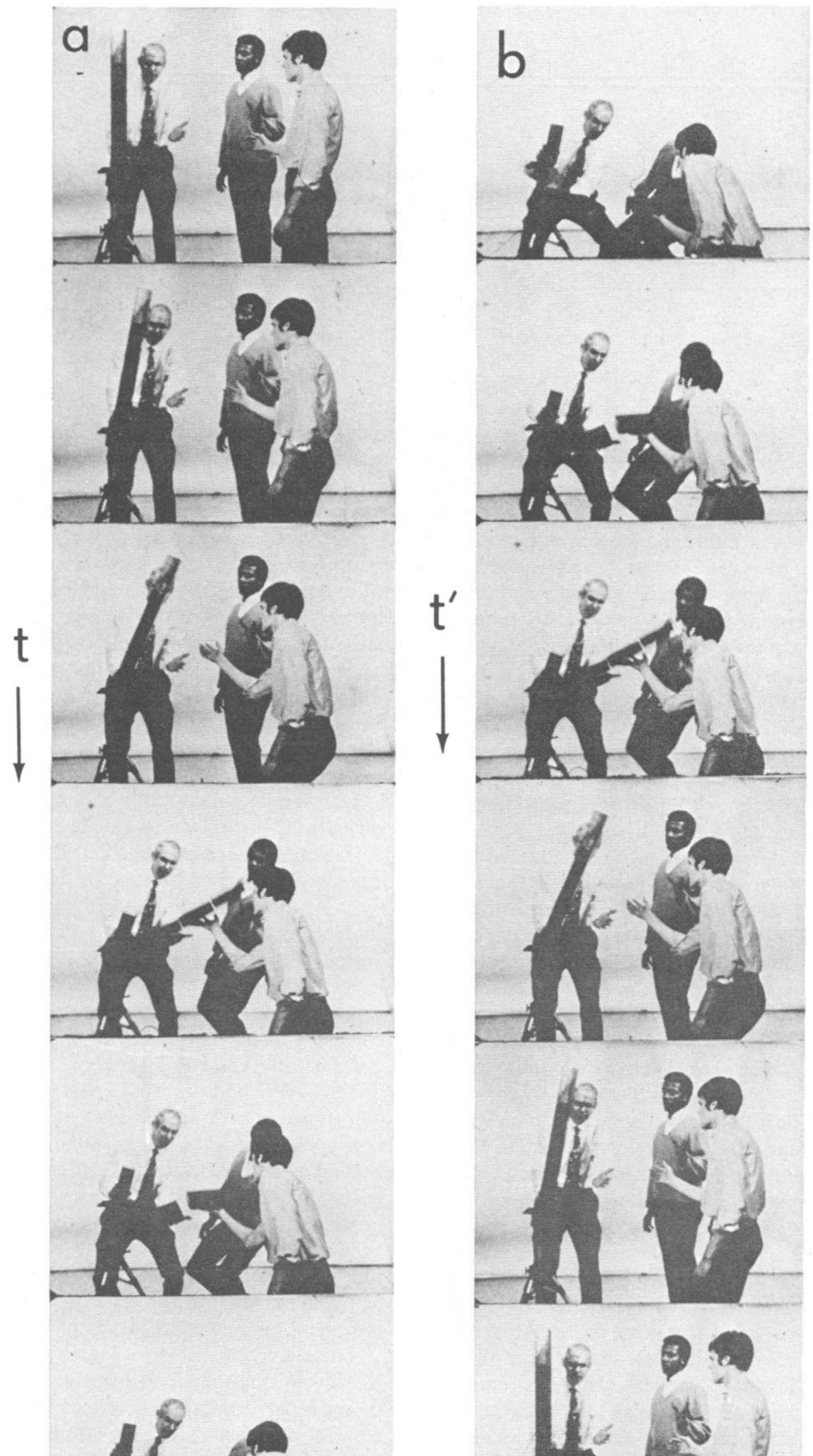


Fig. 6. (a) Motion picture with forward-moving time sequence (t). (b) Same motion picture with reversed time sequence (t').

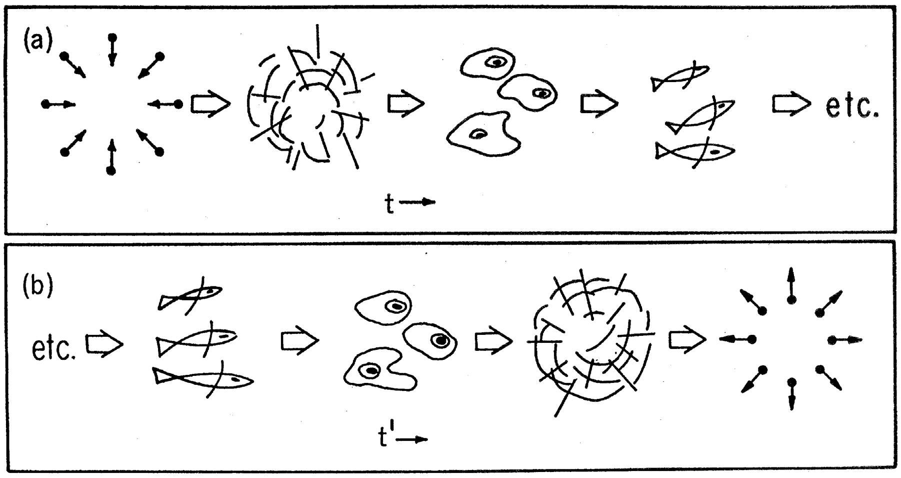


Fig. 7. Model of the evolution of the universe (a) in forward-moving time sequence (t), (b) in reversed motion (t').

This distinction between what is seen in the mirror and what happens in the laboratory under conditions duplicating those in the mirror establishes unambiguously that the distinction between right- and left-handed coordinate systems is absolute, that it is not merely a matter of convention. In fact, the experiment could be used to provide a method for defining a right-handed coordinate system to a being on a remote planet if he knew nothing about the meaning of the words left and right, nor even knew what hand means. He would only need to understand physics and have the appropriate apparatus at his disposal.

Experiments of many different kinds involving weak interactions have confirmed this result. Therefore, the simple notion that there is no distinction between left- and right-handedness is not correct. It is natural to ask whether there is any principle of almost equal simplicity to replace it.

The answer to that question has been formulated in terms of the operation of charge conjugation mentioned earlier. Charge conjugation interchanges all particles with their antiparticles, which have the opposite sign of electric charge. Until the results of the parity experiments were known, it had been assumed by physicists that the laws of motion of antiparticles must be exactly the same as those of the original particles, that is, that they are invariant under charge conjugation. However, it is easy to see that the ^{60}Co experiment just described violates this principle of C-invariance.

If the ^{60}Co nucleus were replaced by the anti-^{60}Co nucleus it would be expected to β-decay by emission of positrons rather than electrons, but the transformation would not be expected to alter the spatial distribution of the particles. Thus, the distribution of β-particles would remain the same as in Fig. 8. On the other hand, when the loop of wire and the rest of the apparatus are replaced by antiparticles, the current in the loop reverses sign (because the current-carrying electrons are replaced by positrons) and the direction of the magnetic field is reversed. Therefore, we again have a contradiction; the mechanism responsible for the unsymmetrical β-decay is different (in fact, opposite) for antiparticles and particles; C-invariance is violated.

However, we now see that if this argument is carried a step further, a form of symmetry is restored. If the charge-conjugation operation is carried out on the image in Fig. 8 rather than in the laboratory, the direction of the magnetic field is restored to the original and the phenomenon looks the same as it did before. The two operations, space reversal (P) and charge conjugation (C), apparently leave the laws of motion unchanged, in which case they can be said to be CP-invariant. This invariance has been verified by many experiments on β-decay, and it was generally believed to be the appropriate substitute for mirror symmetry until the experiment of Christenson *et al.* (2) led to a new contradiction.

The concept of CP-invariance brings us back to the matter at hand because there is a connection between time reversal (T) and CP, known as the CPT-theorem. The theorem states that all physical laws must be invariant under the three transformations simultaneously: If the space coordinates and time variable are reversed and all particles are replaced by antiparticles, the physical laws will remain unchanged. Of course, the proof of such a statement, which is strictly a mathematical theorem, must be based on some assumptions about the mathematical nature of physical laws and on specific hypotheses. The assumptions and hypotheses are so general that it has been very difficult, if not impossible, for physicists to conceive of a reasonable physical theory that violates any of them. However, this situation is a measure of the limitations of theoretical physicists, not of the physical universe, and the theorem itself should be subjected to experimental test.

If the theorem is valid, the invariance of the laws of motion under CP implies their invariance under T, and also a violation of CP-invariance implies a violation of T-invariance. Even if the theorem is disregarded, a demonstrated violation of CP-invariance is an indication that the issue of time reversal must be investigated in every possible way.

Because of the general nature of the assumptions underlying the CPT-theorem, assumptions about the nature of our concepts of space and time and the ultimate localizability of space-time measurements, the tests of CP-, T-, and CPT-invariance go to the very root of our understanding of physical theories. In spite of the highly specialized nature of the phenomena that have, up to this time, demonstrated that there are open questions in this regard, it may turn out that these phenomena are subtle manifestations of an entirely new aspect of physical laws and therefore have an importance to physics that goes far beyond the particular process that is observed.

Nuclear β-decay can be used to test T-invariance directly without going through the intermediary of the CPT-theorem. The nature of such an experiment can be visualized by making use of the concept of reversed motion pictures, as described in connection with Figs. 4 and 5. The β-decay of a polarized neutron offers an excellent example for treatment in this manner. A free neutron undergoes decay into a proton, electron, and neutrino. The polarization of the neutron refers to the orientation of its axis of spin, which can take on only two opposite positions because the quantized spin is ½ in the usual quantum units. Measurements of the neutron polarization and velocities of the products of the decay might result in the situation shown in Fig. 9a, where the neutron is assumed to be at rest before decaying. With enough imagination, we can envision a motion picture of the neutron before decay, showing it as a spinning particle as in Fig. 9a, and we can imagine a separate motion picture of each of the particles produced in the decay, with the velocities indicated in Fig. 9a. Then the reversed motion that would be seen in reversed movies would yield velocities as shown in Fig. 9b. The relative velocities of all particles would be exactly the same as they were in Fig. 9a (rotate by 180° about the spin axis to see this), but the spin is reversed. (Note that only the spin and the velocities, not the decay process itself, are being reversed here. The reversal of decay is taken up later.)

Now, to follow the procedure established in connection with Fig. 4, we repeat the experiment with conditions corresponding to Fig. 9b, that is, with the spin reversed. Invariance under time reversal tells us that the results must be the same as those shown in the reversed movie, namely, the decay characteristics must not depend on whether the spin is pointing up or down. In an actual experiment, this would mean that the intensity of decay for a given configuration of the velocity vectors of the emitted particles does not depend on the component of neutron polarization perpendicular to the decay plane. Experiments on neutron decay have shown no such dependence; therefore, they give no evidence for violation of T-invariance in β-decay (*7*). In the spirit of the CPT-theorem this result is consistent with the evidence for CP-invariance in β-decay.

I must point out that the preceding discussion is somewhat simplified by omission of some of the subtleties associated with the quantum mechanical treatment of the problem. Although the argument is qualitatively correct, the necessity to gloss over the subtleties is unfortunate since all modern work on time reversal is based on Wigner's classic paper on the role of time reversal in quantum mechanics (*8*).

Neutral K-Meson Interferometry

The first clear indication that T-invariance may be violated came about through the failure of CP-invariance in a neutral K-meson phenomenon. Therefore, it is necessary to describe some of the relevant properties of neutral K-mesons and to try to indicate why it is that they provide an especially sensitive detector of these effects.

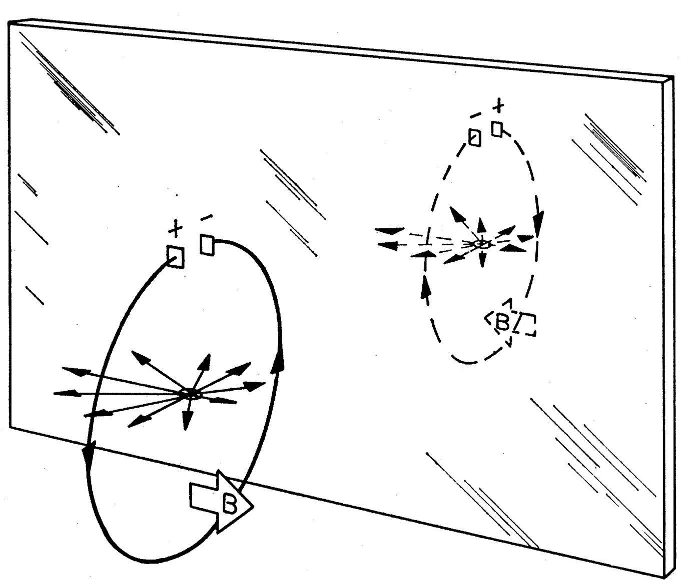
The K-mesons fall into a class of particles called strange because they can be produced only in association with other strange particles in collisions of ordinary particles such as protons, neutrons, or pions. The mass of the K-meson is about one-half the mass of the proton; therefore, energies of billions of electron volts are required to produce them. The K^0, or neutral K-meson, and its antiparticle, the K^0, are both unstable due to their weak interactions. They undergo β-decay (into a π-meson, neutrino, and electron or μ-meson), but they also decay into two or three π-mesons. The 2π-decay mode is the dominant one by a factor of 500 or so.

It was pointed out by Gell-Mann and Pais (*9*) that the availability of the same decay mode to both the K^0 and K^0 would lead to some interesting effects because it means that they are coupled together. No matter how slight this coupling between K^0 and K^0 may be, it has a profound effect because their masses are exactly the same. The situation is similar to one familiar in everyday life: Two harmonic oscillators of exactly the same frequency are profoundly affected if there is the slightest coupling between them; energy introduced into one of the oscillators is completely transferred into the other. Only the time scale for the transfer is determined by the coupling, not the amount of energy transferred.

The effect is even more striking for the K^0-mesons since the observed time-dependent behavior is governed primarily by decay rather than oscillation. Normally, radioactive decay is described by an exponential curve; the number of decaying particles at time t is given by

$$N(t) = N_0 e^{-\Gamma t} \qquad (3)$$

if the number is N_0 at time $t = 0$. The quantity Γ is then the decay rate, which is the reciprocal of the lifetime. When this decay process is described in quantum mechanical terms, it is necessary



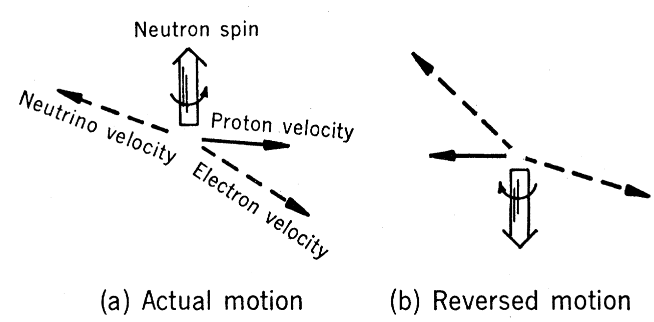


Fig. 8 (left). Space reversal in β-decay. The violation of left-right symmetry is illustrated by the distribution of β-rays (electrons) emitted by ^{60}Co in a magnetic field and its mirror image. The wire loop perpendicular to the mirror produces a magnetic field B parallel to the mirror. The lengths of the arrows illustrate the intensity distribution of β-rays in a plane perpendicular to both the mirror and the wire loop. Fig. 9 (right). Neutron β-decay. Illustration of the way in which measurements of neutron spin polarization and the direction of the velocities of its β-decay products provide a test of time-reversal invariance. (a) Spin and velocity configurations in the laboratory, and (b) spin and velocity configurations as determined from a reversed movie.

to introduce a decay amplitude $A(t)$, and the probability of decay or intensity of the decay mode is obtained by taking its square. (Actually, it is a complex number and the absolute square gives the probability.) The decay amplitude would have the behavior

$$A(t) \sim e^{-\Gamma t/2} \quad (4)$$

and the square would then have the same form as Eq. 3.

The amplitude determined from quantum mechanics also contains an oscillatory factor with a period of oscillation determined by the mass of the particle. For the mass of the K-meson this period is about 10^{-24} second, which means that the amplitude executes some 10^{14} oscillations during the decay lifetime of the particle, an effect that is not easily observed. However, interference between two such oscillations is observable if they are tuned closely, which is the situation with which we shall deal.

Because they are two degenerate (equal mass) particles with a slight coupling between them, the K^0 and $\bar{K}^0$ do not follow the usual exponential-decay behavior of radioactive systems. The amplitude of any decay mode behaves as the sum of two new amplitudes having distinctly different characteristic lifetimes. In fact, as we shall see later, only one of these amplitudes would be associated with decay into the dominant 2π-decay mode if CP-invariance were valid; the other amplitude would be associated with decay into the modes that are 500 times rarer.

These new amplitudes may be ascribed to two new K-mesons, one, called K_S, having a short lifetime (about 10^{-10} second) and the other, called K_L, having a lifetime 500 times longer. The amplitudes, $A_S(t)$ and $A_L(t)$, behave in the normal fashion as very rapidly oscillating quantities damped slowly by the decay. There is a slight difference in the masses of K_S and K_L caused by the weak coupling between K^0 and $\bar{K}^0$ leading to a slight frequency difference between the oscillations of the two amplitudes (this would correspond to the frequency shift of coupled harmonic oscillators). This frequency difference is comparable in magnitude to the decay rate, Γ_S, for the fast decay since it is caused by the same weak coupling.

The connection with the behavior of K^0-mesons and $\bar{K}^0$-mesons is as follows: If K^0 (or $\bar{K}^0$) mesons are produced at a particular instant of time,

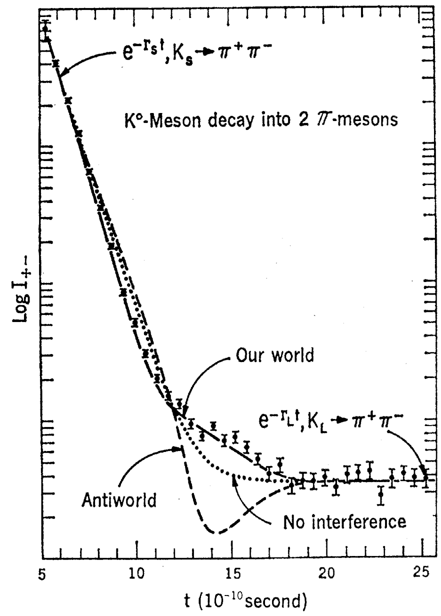


Fig. 10. Intensity of the 2π-decay mode of the K^0-meson as a function of time. (Solid curve) A fit to the data in the real world. (Dashed curve) Appearance to be expected in the antiworld. (Dotted curve) Appearance to be expected for two independent, noninterfering decay processes.

the time dependence of their decay is governed by an amplitude composed additively of the amplitudes $A_S(t)$ and $A_L(t)$. Since the intensity of decay into a particular mode is given by the square of the amplitude, the situation is similar to that in any wave phenomenon where two coherent waves are combined; there is interference. In this case, the interference pattern is dominated by the decay; the intensity is the sum of three terms, and two of them represent normal exponential decay of the K_S and K_L. The third term describes the interference, and it has its greatest effect on the intensity when the K_S has decayed to such a point that its intensity is comparable to the intensity of the K_L. This can be seen in Fig. 10, where the decay of a K^0 beam into two π-mesons (π^+ and π^-) is shown by the solid curve and data points (*10*). The dotted curve represents the sum of the two noninterfering terms.

The interference term depends on the small difference in oscillation frequency of the two amplitudes, and it also depends on the relative phase of the two waves. It is possible to produce various mixtures of the K_S and K_L by several experimental techniques and, therefore, to vary the interference pattern. Consequently, both the frequency difference and the phase can be measured (*11*). The frequency difference is a measure of the difference in mass between the K_S and K_L, which has been obtained in this way to within several percent. The fact that one can measure this mass difference, which is 10^{-14} of the total mass, is an indication of the sensitivity of K^0-meson interferometry.

Violation of CP-Invariance

The enormous difference in lifetime between the K_S and K_L makes it possible to produce a beam of pure K_L-mesons. The lifetime of K_S-mesons is approximately 10^{-10} second, so that even at velocities close to the speed of light their numbers are substantially depleted in a distance of a few centimeters. Therefore, a beam initially consisting of a large number of K^0 (or $\bar{K}^0$) mesons becomes an almost pure beam of K_L after traversing a distance of several meters. In this way, a number of important experiments have been carried out on the properties of K_L-mesons, including the experiment that led to the discovery that CP-invariance is violated. The conclusion that CP-invariance is violated followed from the discovery that the long-lived K_L decays into two π-mesons (*2*), a conclusion that is a straightforward consequence of the theory of K^0-meson phenomena, although the connection between the observation and the conclusion is by no means obvious. It is possible to visualize this connection by means of imaginary experiments similar to those used here for understanding the consequences of space reversal, phantasmagorias having much in common with Alice's experiences on her trip through the looking glass. An updated version of the looking glass is needed here, one that not only reverses space (P) but also replaces every particle by its antiparticle (C). I shall refer to the environment inside this looking glass as the antiworld.

The decay mode $K_L \rightarrow \pi^+\pi^-$ observed by Christenson *et al.* (*2*) has the same appearance in the world and antiworld. In order to establish a connection with CP-violation more information is needed, in particular that there is coherence between the K_S and K_L leading to interference between their decay modes. The existence of coherence was established by Fitch *et al.* (*12*), thereby establishing unambiguously the violation of CP.

The interference in the 2π-decay mode has already been described in connection with Fig. 10, and this phenomenon may now be used to demonstrate that the existence of a $K_L \rightarrow 2\pi$ mode does indeed imply the violation of CP. It is only necessary to investigate the appearance of the same phenomenon in the antiworld. For that purpose, we need to know how the interference term will behave in going from the world to the antiworld, since the other terms describing normal radioactive decay certainly will not be affected by going through the looking glass.

The interference term depends on the product $A_S A_L$ of the amplitudes for decay of the K_S and K_L and on their relative phase. Both A_S and A_L are linear combinations of the amplitudes A and $\bar{A}$ for decay of the K^0 and $\bar{K}^0$ with coefficients depending on the way in which K^0 and $\bar{K}^0$ are coupled to one another. The act of going from the world into the antiworld interchanges the roles of K^0 and $\bar{K}^0$, and it can be shown on very general grounds that this change leaves A_S unchanged and merely changes the sign of A_L. Therefore, the interference term has the opposite sign in the antiworld, and the decay curve will have the appearance of the dashed curve in Fig. 10 when seen through the looking glass. Therefore, the existence of the interference means that the law of decay is different when viewed in the world and antiworld, a clear violation of CP-invariance (*13*). You will note that this observation does not entail the identification of the sign of an electric charge and is therefore not concerned with any convention as to the definition of particle and antiparticle. Only the measurement of an intensity as a function of time is needed, and the experiment could, in fact, be used to define the distinction between particles and antiparticles in an absolute sense.

The magnitude of A_L compared to A_S is quite small, about 2×10^{-3}, as can be seen from Fig. 10. This means that the CP-violating effects, as observed in this phenomenon, are quite small. The dominant decay amplitude is the part of A_S that would occur if CP-invariance were valid, any corrections to A_S associated with CP-violation being of the same magnitude as A_L. The small size of the effect is the reason that the long-lived and short-lived K-mesons are so distinct in spite of the CP-violation. At the same time, it is the small size of the effect that is a cause for mystification; the breaking of a symmetry should occur with a resounding roar rather than with a whimper. It may be (in fact, it seems likely to me) that this mystery is due to the way the phenomenon is observed, that the direct interaction responsible for the symmetry breaking is large but only a very small indirect effect is being seen by means of the sensitive K^0 interferometry.

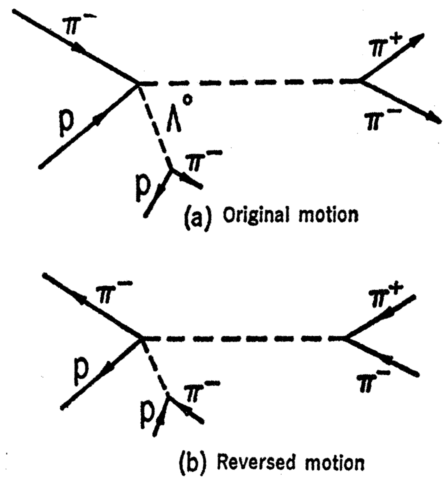


Fig. 11. Illustration of motion reversal in the K^0-meson production and decay process (a) as seen in the real world, and (b) as seen in a reversed movie.

Confirmation of CP-violation has been found in the charge asymmetry experiment, in which the relative rates of decay are measured for the modes $K_L \rightarrow \pi^- l^+ \nu$ and $K_L \rightarrow \pi^+ l^- \bar{\nu}$, where l (for lepton) stands for either the electron (positron) or μ-meson, and ν and $\bar{\nu}$ stand for the neutrino and antineutrino, respectively. A difference in these rates is reversed in going from the world to the antiworld, but the K_L is unchanged by the transformation. Such a difference has been observed (*14*), confirming the existence of an absolute distinction between world and antiworld, that is, CP-violation. The magnitude of the effect is again a few tenths of a percent.

Violation of T-Invariance

The violation of CP-invariance implies that T-invariance is also violated if the CPT-theorem is taken to be a law of nature. However, my purpose is to provide a visual understanding of the connection between time reversal and the experiments, and that connection is certainly not obvious in the case of the 2π-decay mode of the long-lived neutral K-meson. In fact, the application of time reversal to any decaying system can cause confusion because the direct introduction of $t' = -t$ in an exponential decay leads to an unlimited growing exponential, which is not a credible behavior. As usual, this credibility problem has to do with initial conditions, and one must look at the entire process beginning with the production of the neutral K-mesons and ending with the decay in order to understand the effect of time reversal.

Figure 11a shows a typical production and decay process for K^0-mesons. A target of protons is bombarded with π-mesons to produce a K^0-meson in association with another unstable neutral particle known as the Λ-hyperon, which decays into a proton and a π^--meson. The K^0 then moves on to eventually decay into two π-mesons. When reversed, a motion picture of this process would show the reaction illustrated in Fig. 11b: Two π-mesons would impinge upon one another at one instant, and at some other instant, at another point in space, a proton and a π^--meson would enter the picture and collide. Later, at a third point in space, an outgoing π^--meson and a proton at rest would materalize.

The test procedure to be used here is the one proposed in connection with Fig. 4 or Fig. 5. It involves carrying out, in the laboratory, this reversed experiment, a formidable task (*15*). However, we need only imagine the experiment and imagine comparing the result with the reversed movie to see whether the quantitative behavior is the same or not. The comparison is more easily visualized by imagining that the procedure is turned around: The imaginary reversed laboratory experiment is the one that is photographed, and the reversal of this movie is compared with the original laboratory experiment.

The advantage of this last procedure is that the reversed movie shows an experiment on the t' scale that exactly satisfies the initial conditions of the original experiment, namely, production of a K^0-meson followed (on the t' scale) by its decay. The question is then how the results of this experiment on the t' scale would compare with the original decay curve on the t scale, shown in Fig. 10.

First, a theory of this production and decay process (on the t' scale) will lead to the normal exponentially decaying terms. What the theory has to say about the interference term is determined by the properties of the product $A_S A_L$ of the decay amplitudes of the K_S and K_L, in this case, their time-reversal properties.

The amplitude A_S is dominated by the "natural" decay into the 2π mode since it includes the preponderant CP-invariant amplitude. Any effect of time reversal would appear in the small correction terms in A_S due to CP-violation; therefore, it is a good approximation to treat A_S as being unchanged under time reversal. The time-reversal properties of the interference term $A_S \cdot A_L$ then depend on the behavior of A_L.

An analysis of the behavior of A_L under Wigner time reversal (T), taking into account the way the initial conditions must be set in going from the t scale to the t' scale, leads to the conclusion that, in general, A_L is the sum of two terms

$$A_L = A_L^+ + A_L^- \qquad (5)$$

such that A_L^+ is unchanged under T and A_L^- changes sign under T. Therefore, in the general case, one may expect that on the t' scale the curve for the decay of a K^0-meson into two π-mesons might be different from either the solid or the dashed curve in Fig. 10. However, if $A_L^+ = 0$, as one would expect on the basis of the CPT-theorem, the dashed curve would be obtained, the same curve that is found in the antiworld, and this would make it possible, in principle, to distinguish a time-reversed world from the original world. On the other hand, if $A_L^- = 0$, the solid curve would be obtained, T-invariance would be valid, the time-reversed world could not be distinguished, but, because an antiworld could be distinguished, the CPT-theorem would be violated.

Since the transformation of Fig. 10 described here concerns a thought-experiment that cannot be realized in fact, it is only an exercise telling us what we must know to establish the consequences of time reversal. The message to be taken from the exercise is that $A_L \neq 0$ is not enough information to establish a violation of T-invariance although it does establish a violation of CP-invariance. The necessary condition is that $A_L^- \neq 0$.

The question then is whether there is any way to separate the contributions of A_L^+ and A_L^- to A_L. It can be shown by means of the general theory of the K^0 phenomena, without explicit use of the CPT-theorem, that the phase associated with A_L should be different if $A_L^+ = 0$ from its value if $A_L^- = 0$, and the phase can be estimated in each case from detailed experimental data on the relative rates of various decay modes, including not only the $\pi^+\pi^-$ mode discussed here but also the $\pi^0\pi^0$ mode and all other modes. Therefore, a determination of the phase associated with A_L can be used to separate the contributions of A_L^+ and A_L^-. This phase has been determined experimentally by K^0 interferometry. The result of analysis of many experiments is that the phase associated with A_L is consistent with the phase expected if $A_L^+ = 0$ and inconsistent with that expected if $A_L^- = 0$ (*16*), which shows that $A_L^- \neq 0$ and, therefore, establishes the breakdown of the principle of universal time-reversal invariance.

Although this result also shows that A_L^+ must be smaller than A_L^-, the accuracy with which the measurements have been made does not place a precise limit on the magnitude of A_L^+; therefore, it does not provide precise information concerning the limits of validity of the CPT-theorem, which would require that $A_L^+ = 0$.

Conclusion

The breaking of time-reversal symmetry which is observed occurs as a tiny effect even as measured against the scale of the weak interactions. Nevertheless, this violation of what has appeared to be a natural and universal symmetry must have deep implications (*17*). It certainly would be helpful to have a more direct measure of the violation of T-invariance, such as one that could be associated with angular correlations between the spin and velocity vectors of decay products, as described in connection with Fig. 9. All measurements of this kind made up to this time on the decay of strange particles as well as on the decay of ordinary nuclear particles are consistent with T-invariance. Additional and more precise measurements must be made in order to match the precision achieved with K^0-meson interferometry, but it is clear that the effects of violation of T-invariance in weak decay processes are very small.

The limitations of the decay measurements for understanding a deep process are severe. The variables are limited by the number of particles undergoing weak decay. In particular, the range of energies over which the measurements can be made is limited (by the relation $E = mc^2$) to the mass changes that take place in the decay, which are of the order of several hundred million electron volts (Mev). Possibly, the interactions violating CP-invariance and T-invariance are large, that is, comparable to the full weak interaction at higher energies, and the reason for our observing only very small effects is that in decay phenomena we see only a residual low-energy "tail" on a process at very high energy. This would be the case, for example, if the breaking of the symmetry occurs only over very small distances and time intervals, since the space-time intervals of importance in a physical observation correspond to the wavelengths associated with the particles being used, and these wavelengths are small in inverse proportion to the energy, at high energy. (The distance corresponding to 100 Mev is about 2×10^{-13} centimeter or, roughly, the size of atomic nuclei.)

In order to avoid the energy limitations of decay processes, we may investigate weak interactions by means of neutrino beams, the only known interaction of the neutrino being due to the weak coupling. Production of neutrino beams at high energy requires the use of particle accelerators at 100 or more billion electron volts (Gev). Laboratory energies of this magnitude (200- to 400-Gev protons) will soon be attained for the first time at the National Accelerator Laboratory in Batavia, Illinois, which will provide the first opportunity to determine whether there is a connection between the nature of measurements over small space-time intervals and the existence of a violation of the time-reversal principle. Therefore, it can be hoped that within the next few years both experiments at very high energies and experiments with high precision on the many available decay phenomena will shed light on the source of this mysterious violation of the time-reversal symmetry.

Reference and Notes

1. R. G. Sachs, *Science* **140**, 1284 (1963).
2. J. H. Christenson, J. W. Cronin, V. L. Fitch, R. Turlay, *Phys. Rev. Lett.* **13**, 138 (1964); *Phys. Rev. B* **140**, 74 (1965).
3. T. D. Lee and C. N. Yang, *Phys. Rev.* **104**, 254 (1956); C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, R. P. Hudson, *ibid.* **105**, 1413 (1957).
4. The term "calendar" rather than the more usual "clock" is used advisedly here since a clock is generally used to provide a measure of time intervals rather than the absolute measure of time necessary for considerations of time reversal.
5. J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (1902), reprinted in *The Collected Works of J. Willard Gibbs* (Longmans, Green, New York, 1931), p. 144.

6. A demonstration that this is indeed possible can be given by means of an apparatus created by C. Hutchison. This demonstration is shown on film No. 991135, entitled "Time Reversal," produced by Bell Telephone Laboratories, Murray Hill, New Jersey. It is a beautiful illustration of Gibbs' point, but its real bearing on the time-reversal question is related only to the matter of coarse-grained and fine-grained information. The irrelevance of time reversal per se has to do with the absence of any dynamics. In fact, the trick in making this demonstration work is to avoid dynamical effects by using a very viscous fluid and stirring very slowly.
7. M. T. Burgy, V. E. Krohn, T. B. Novey, G. R. Ringo, V. L. Telegdi, *Phys. Rev. Lett.* **1**, 324 (1958); B. G. Erozolimskii, L. N. Bonderenko, Y. A. Mostovoi, B. A. Obinyakov, V. P. Zakharova, V. A. Titov, *Sov. J. Nucl. Phys.* **11**, 583 (1970).
8. E. P. Wigner, *Nachr. Ges. Wiss. Göttingen* (No. 32) (1932), p. 35.
9. M. Gell-Mann and A. Pais, *Phys. Rev.* **97**, 1387 (1955).
10. This graph was kindly provided by V. Telegdi. The data points are based on the data of D. A. Jensen, S. H. Aronson, R. D. Ehrlich, D. Fryberger, C. Nissim-Sabat, V. L. Telegdi, H. Goldberg, and J. Solomon [*Phys. Rev. Lett.* **23**, 615 (1969)].
11. The most recent and highest-precision measurements of this kind are those reported in S. H. Aronson, R. D. Ehrlich, H. Hofer, D. A. Jensen, R. A. Swanson, V. L. Telegdi, H. Goldberg, J. Solomon, D. Fryberger, *ibid.* **25**, 1057 (1970); M. Cullen, P. Darriulat, J. Deutsch, H. Foeth, G. Groves, M. Holder, K. Kleinknecht, E. Rademacher, C. Rubbia, D. Shambroom, M. Scire, A. Stande, K. Tittel, *Phys. Lett. B* **32**, 523 (1970); R. K. Carnegie, R. Cester, V. L. Fitch, M. Strovink, L. R. Sulak, *Phys. Rev. D* **4**, 1 (1971).
12. V. L. Fitch, R. F. Roth, J. S. Russ, W. Vernon, *Phys. Rev. Lett.* **15**, 73 (1965).
13. This way of approaching a description of the violation of CP-invariance was first suggested to me by V. L. Telegdi, who then raised the question: Is there any corresponding way to show explicitly that the 2π mode violates T-invariance? The answer is given here.
14. S. Bennett, D. Nygren, H. Saal, J. Steinberger, J. Sunderland, *Phys. Rev. Lett.* **19**, 993 (1967); D. Dorfan, J. Enstrom, D. Raymond, M. Schwartz, S. Wojcicki, D. H. Miller, M. Paciotti, *ibid.*, p. 989; J. Marx, D. Nygren, J. Peoples, T. Kirk, J. Steinberger, *Phys. Lett. B* **32**, 219 (1970).
15. F. Coester, *Phys. Rev.* **84**, 1259 (1951).
16. R. C. Casella, *Phys. Rev. Lett.* **21**, 1128 (1968); *ibid.* **22**, 554 (1969).
17. Whenever the word deep is used in connection with physics it evokes a quotation from Neils Bohr that has been conveyed by word of mouth among physicists: "In order to define a deep statement it is first necessary to define a clear statement. A clear statement is one to which the contrary statement is either true or false. A deep statement is a statement to which the contrary is another deep statement."
18. Part of the work reported here was supported by the U.S. Atomic Energy Commission.

Energetics and Pollination Ecology

The energetics of pollinators may have wide implications in floral biology and community ecology.

Bernd Heinrich and Peter H. Raven

Color, shape, and odor are well-known characteristics of flowers which partly determine the kinds of animal pollinators that visit them (*1–4*). In turn, these characteristics are influenced, in an evolutionary sense, by the activities of the pollinators. The plants and the pollinators are part of a dynamic, coevolving system, of which the features mentioned above are only a part. Little attention has been paid to another very significant parameter of the system: the caloric reward (*5*) provided by the flowers of particular plant species. The energy budget of pollinators in relation to the food reward provided by the flowers they visit has been investigated in a few instances (for example, *6–10*), but the role that this energy budget plays in the evolution of flowering plants has not been discussed previously.

Dr. Heinrich is assistant professor of entomology at the University of California, Berkeley; Dr. Raven is the director of the Missouri Botanical Garden and professor of biology, Washington University, St. Louis.

In this article we aim to provide a synthesis and point of reference for this aspect of evolution by examining floral biology and the operation of outcrossing systems in plant populations from the standpoint of the energetics of the pollinators. We attempt to point out some potentially unifying theories in this area, hoping that the recognition of these will lead the way to quantitative investigations.

Energy Balance and Cross-Pollination

That animal pollinators usually restrict their visits to the flowers of a particular plant species (*11*) is only one factor promoting outcrossing. For outcrossing to result it is also necessary that the pollinator not confine its visits to a single flower or to the flowers of a single plant. Through evolution plants could presumably reduce the frequency of such repeated visits to the same flower or plant by limiting the caloric reward that is presented at any one time. However, the meaning of a certain caloric reward can be assessed only in relation to a particular animal, because the differences between the energy requirements of different pollinators may be great. For instance, a 100-milligram bumblebee that lands on each flower may expend about 0.08 calorie per minute when walking (*12*), while 3-gram sphinx moths (*13*) and hummingbirds (*14*) expend energy while hovering at a rate of about 11 cal/min, more than a 140-fold difference. However, many other aspects of the biology of the pollinators affect net energy expenditure.

The specific amounts of nectar per flower, in terms of calories of food energy, that would promote maximum cross-pollination, are related to the characteristic rate of energy expenditure of the pollinators. For example, the flowers of saguaro cactus, *Carnegiea gigantea* (Engelm.) Britt. & Rose, which are visited by many different insects as well as by some birds and bats (*15*), produce large quantities of nectar. Some individual honeybees (*Apis mellifica* is not native to North America) tend to limit their visits on numerous foraging trips to specific flowers of the same cactus plant (*16*). Since these plants are self-incompatible, such visits do not result in seed production. Secondly, those honeybees that take the nectar from saguaro may also reduce the subsequent attractiveness of the flowers to birds and bats, which have higher rates of energy expenditure and whose visits will more often result in outcrossing.

Animals with high energy requirements may not forage at the flowers of