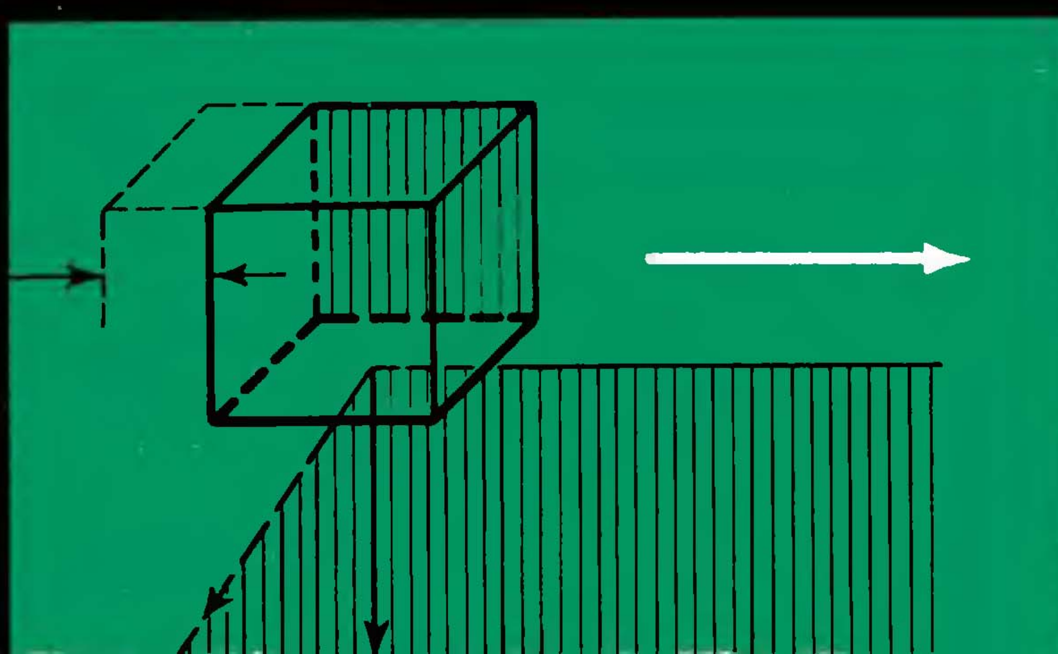
COMMISSION ON COLLEGE PHYSICS

An Introduction to the Special Theory of Relativity

ROBERT KATZ



An Introduction to the SPECIAL THEORY OF RELATIVITY

ROBERT KATZ
Professor of Physics
Kansas State University

An Introduction to the SPECIAL THEORY OF RELATIVITY

Published for
The Commission on College Physics

FOR STEVE AND JOHN

Preface

It is the purpose of this book to provide an introduction to the Special Theory of Relativity which is accessible to any student who has had an introduction to general physics and some slight acquaintance with the calculus. Much of the material is at a level suitable for high school students who have had advanced placement in physics and mathematics. Since some of the exposition, particularly part of Chapter 6, is presented here in book form for the first time, even terminal graduate students in physics may find the material profitable.

The subject matter of the book has been taught, in the form presented here, to first-year graduate students at Kansas State University for many years. Much of it has sifted into courses at lower level.

The typescript was read by Professors E. U. Condon, University of Colorado, and Melba Phillips, University of Chicago. I wish to thank them for their helpful criticism and valuable suggestions.

May 26, 1964 ROBERT KATZ
Manhattan, Kansas

Table of Contents

Preface 5

1 The Galilean Transformation 9

The Inertial Frame, 9; The Galilean Transformation, 11; The Speed of Light, 17; The Special Theory of Relativity, 24; The Rod Clock, 25

2 The Lorentz Transformation 29

The Lorentz Transformation, 29; Simultaneity and Time Sequence, 31; Time Dilation, 33; Lorentz Contraction, 34; Velocity Transformations, 37; The Fizeau Experiment, 38; Aberration, 40; Visual Appearance of Rapidly Moving Objects, 44; Transformation of Acceleration, 48

3 Force and Motion 51

Introduction, 51; Newton's Second Law, 53; The Equivalence Principle, 56; Transformation of Forces, 58

4 Energy and Momentum 61

Work, 61; Kinetic Energy (Nonrelativistic), 63; Kinetic Energy (Relativistic), 64; Conservation of Linear Momentum, 68; Center of Mass (Nonrelativistic), 71; Transformation of Momentum and Energy (Relativistic), 75; Center of Mass (Relativistic), 78

5 Some Relativistic Phenomena 83

Pair Production and Annihilation, 83; Compton Effect, 85; Threshold for π_0 Meson Production, 87; Doppler Effect, 89; Relativistic Dynamics, 92; Gravitation and Light, 96

6 Relativity and Electromagnetism 103

Introduction, 103; The Lorentz Force, 105; Magnetization and Polarization, 106; Transformations of Fields and Flux Densities, 108; Electromagnetic Induction, 110; Field of a Moving Charge, 114; Transformation of Polarization and Magnetization, 117; The Unipolar Generator, 121; Postscript, 123

General Bibliography 127

Index 129

1 *The Galilean Transformation*

§1-1 The Inertial Frame Does the earth move around the sun or the sun around the earth? Do we describe the motion of the moon by saying that it traces a nearly circular path around the earth, or a somewhat more complicated path around the sun? These questions are associated with the choice of a frame of reference, a background against which we imagine motions to be referred. The choice is not a simple one, since it has philosophic as well as physical overtones. The central position accorded to man in the universe led to the belief that the sun and stars circulate about the earth, and in consequence, the universe of Ptolemy (A.D. 127-151) demanded that heavenly bodies move in paths compounded from circles, called epicycles. The path of a point which moves on a small circle, whose center in turn moves on a larger circle, is called an epicycle. In part the composition of the paths of heavenly bodies from circles arose from the view that only the most perfect figure, the circle, was appropriate to any object in the heavens. Copernicus' (1473-1543) relocation of the origin of the coordinate frame to the sun led to a new grouping of the earth and planets as a system of solar satellites and resulted in a new set of physical laws, whose truth was universal, valid on earth and in heaven. These were the laws of Newton (1642-1727).

The philosophic implications of the relocation of the origin of astronomical coordinates have been widely elaborated under names like rationalism, or naturalism, or secularism. It is our purpose here to emphasize the importance of the choice of a coordinate frame in physics, though clearly the choice has far broader implications in the affairs of men.

The logical choice of a coordinate frame to Newton was one in which the stars were at rest. There are some physicists today to whom this choice seems a proper one. A concept called *Mach's Principle,* propounded by Ernst Mach (1838-1916), states that inertia is entirely due to the mutual action of matter, so that force is required to accelerate matter only because it is displaced relative to the other matter of the universe. More generally, physicists place a weaker restriction on the choice of a coordinate frame: the frame must be inertial. *An inertial frame is one in which Newton's first law is true.*

For many centuries before Newton, inquiries into the way the world was structured led to the belief that the natural state of things on earth was a state of rest—objects left alone remained at rest. To set them in motion and to keep them in motion was thought to require the application of a force. Newton's formulation gave equal importance to rest and to uniform motion in a straight line, and asserted that either of these states was natural; that is, either of these conditions would continue indefinitely if no force was applied. This was in great contradiction to earlier beliefs which supposed that force was required to sustain motion as well as to initiate motion.

Newton's First Law states: A body at rest remains at rest, and a body in motion will continue in motion with steady speed in a straight line as long as no outside force acts on the body.

Today we accept the first law as axiomatic, and use it as a recipe for finding a coordinate frame in which to describe physical systems. Unless otherwise stated all physical laws stated in this book (and in most others) are with respect to an inertial frame. The symbols which appear in physical equations represent quantities which are usually measured by observers stationed in an inertial frame.

A test for an inertial frame is simplicity itself. A frame of reference may be thought of as a mesh of lines. In free space, in the absence of gravitational or other force fields, a particle set down in an inertial frame may always be found there. If the particle is set in motion, it will move with steady speed in a straight line. Its coordinates, measured in this frame, will satisfy the equation of a straight line, and its motion along this line will be with constant speed. If the particle does not remain at the point at which it is placed, or does not move in a straight line with steady speed, then the frame is not inertial. A perfectly smooth, level, spinning turntable on earth is not an inertial frame. A particle placed on this turntable will remain at a single point with respect to the earth as the turntable slips under it, but it will describe a circular path with respect to the turntable. Thus the turntable is not an inertial frame, for though no forces act on the particle it is neither at rest nor moving in a straight line with constant speed with respect to the turntable.

Experiment has shown that Newton's frame of reference, fixed in the stars, is an inertial frame. A coordinate frame fixed in the earth is not inertial, due to the earth's rotation about its axis, and about the sun as well. For most practical purposes, and to the accuracy with which many experiments are performed, the fact that a frame fixed in the earth is not inertial does not alter the outcome of experiments sufficiently to cause concern. In fact it was only through the *Foucault pendulum* that the rotation of the earth was first clearly demonstrated, in 1851. Foucault (1819-1868) suspended a heavy pendulum bob by a long flexible wire, and observed that the plane of vibration of the pendulum rotated with respect to the floor. But there was no way the support point could twist the plane in which the pendulum oscillated once it had been started. He concluded that the floor was rotating beneath the pendulum. We infer from such an experiment that the earth is not an inertial frame. There are everyday experiences in which the distinction between inertial and non-inertial frames must be made. One of these lies in the circulation of the atmosphere, and another in the meaning of the term *centrifugal force.*

§1-2 The Galilean Transformation Through his studies of projectile motion, Galileo (1564-1642) concluded that the motion of a projectile launched from the ground at an arbitrary angle could be derived from the motion of a projectile launched straight up. A projectile thrown straight up from a uniformly moving cart would be seen from the cart as moving straight up and down, but the motion seen from the ground could be predicted by superimposing the motion of the cart onto that of the projectile. These ideas of Galileo are the basis of the *Galilean transformation,* through which we relate motions as they are observed in two different inertial frames. We will see that Galilean transformations must be superseded by a different transformation, the Lorentz transformation, when the two frames move at great speeds with respect to each other. But for ordinary experience the Galilean transformation is much simpler and is quite satisfactory.

Let us suppose that two sets of observers are available to study a problem. One set is at rest with respect to the ground, and makes its measurements with respect to a coordinate frame and a set of clocks fixed in the ground and synchronized to read alike. We call the set of observers and the observations they make the *unprimed set,* and refer to their measured coordinates and time as $x, y, z, t,$ without primed superscripts. A second set of observers is at rest with respect to a second coordinate frame, the primed frame, which moves with constant velocity with respect to the unprimed frame, as if the second set of observers were riding on a large space platform. We will call the primed observers and their observations the *primed set.* We will refer to their observations as x', y', z', t'.

We choose the two sets of coordinates to be parallel. The x' axis is parallel to the x axis, the y' axis to the y axis, the z' axis parallel to the z axis. Further, we choose the orientation of the axes such that the motion of the primed frame is parallel to the x axis. The primed frame moves in the $+x$ direction with velocity $\mathbf{V}$ with respect to the unprimed frame. For convenience we choose the origins of the two sets of coordinates to coincide at times $t = t' = 0$, as in Fig. 1-2.1.

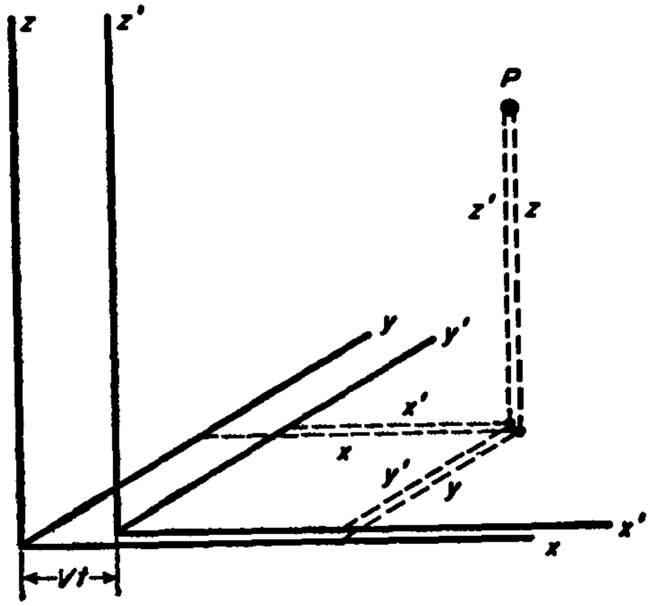


FIG. 1-2.1

Suppose that a particle is located at a point P at a particular time, and that measurements of the coordinates and the time are made by the two sets of observers. The measured values can be related by the equations

$$x' = x - Vt, \quad (1\text{-}2.1a)$$

$$y' = y, \quad (1\text{-}2.1b)$$

$$z' = z, \quad (1\text{-}2.1c)$$

and

$$t' = t. \quad (1\text{-}2.1d)$$

The equations 1-2.1 are the *Galilean transformation* equations, which relate the observations of position and time made by two sets of observers located on two different inertial frames, as described.

The point P may be on the path of a projectile. To describe the motion properly we need to say something about the velocity and the acceleration of the projectile. This implies that we wish to study differential increments of the coordinates in the two frames. We take differentials of the transformation equations 1-2.1, recalling that V is constant, to find

$$dx' = dx - V\,dt, \tag{1-2.2a}$$

$$dy' = dy, \tag{1-2.2b}$$

$$dz' = dz, \tag{1-2.2c}$$

and

$$dt' = dt. \tag{1-2.2d}$$

Now we divide each of the equations 1-2.2a-c by Eq. 1-2.2d to find transformation equations for the components of the velocity. We have

$$\frac{dx'}{dt'} = \frac{dx}{dt} - V, \quad \text{or} \quad v_{x'} = v_x - V. \tag{1-2.3a}$$

$$\frac{dy'}{dt'} = \frac{dy}{dt}, \quad \text{or} \quad v_{y'} = v_y. \tag{1-2.3b}$$

$$\frac{dz'}{dt'} = \frac{dz}{dt}, \quad \text{or} \quad v_{z'} = v_z. \tag{1-2.3c}$$

These results may be put into vector form by multiplying each of these equations by the appropriate unit vectors. A *unit vector* is a dimensionless quantity which indicates direction. Thus the unit vector $\mathbf{1}_x$ is parallel to the x axis and of magnitude 1. Our primed and unprimed coordinate axes are parallel so that $\mathbf{1}_{x'} = \mathbf{1}_x$, and so on. Thus if we multiply Eq. 1-2.3a by $\mathbf{1}_x$, Eq. 1-2.3b by $\mathbf{1}_y$, and Eq. 1-2.3c by $\mathbf{1}_z$ and add these equations, we find that

$$v_{x'}\mathbf{1}_{x'} + v_{y'}\mathbf{1}_{y'} + v_{z'}\mathbf{1}_{z'} = v_x\mathbf{1}_x + v_y\mathbf{1}_y + v_z\mathbf{1}_z - V\mathbf{1}_x,$$

which may be condensed to read

$$\mathbf{v}' = \mathbf{v} - \mathbf{V}. \tag{1-2.4}$$

If we transpose the velocity of the primed frame with respect to the unprimed frame, we obtain

$$\mathbf{v} = \mathbf{v}' + \mathbf{V}. \tag{1-2.5}$$

It is through Eq. 1-2.5 that the airspeed of an airplane is converted to groundspeed, using knowledge of the velocity of the air with respect to the ground. When we speak of windspeed we generally mean the speed an air mass is moving with respect to the ground. When we speak of the airspeed of an airplane, or its cruising speed, we mean the speed with which the airplane moves with respect to the air mass. For navigation with respect to the earth the groundspeed is required, or more precisely, the velocity with respect to the ground is required. The Galilean transformation makes air navigation possible.

In ordinary conversation we are sometimes careless about specifying the frame to which a position or a motion is referred. When we give the cruising speed of an airplane we may be vague in noting that it is with respect to the air mass through which the plane is moving. When we give the windspeed we may be vague in noting that this is with respect to the ground. Consider the question, with what speed is the air blowing past the wings of an airplane whose cruising speed is 200 miles per hour (mph) when it flies into a 50 mph headwind? By the definition of cruising speed the answer is 200 mph, when the airplane is operating properly. If the unprimed coordinate frame is fixed in the ground and the primed set is fixed in the air mass, then the design of the airplane implies that its speed v' is 200 mph for proper operation. If v' is very much less than 200 mph, the airplane falls down; if v' is much greater than 200 mph, the wings fall off.

Neither the magnitude of the velocity nor the direction of the path (the direction of the velocity vector) is the same in the two frames. As a simple illustration consider the motion of a boat on a river. Suppose that the water moves east at 3 mph and that a boat heads north at 4 mph with respect to the water (see Fig. 1-2.2). If we take east as the positive x direction and north as the positive y direction, then the motion of the boat with respect to the primed frame (fixed in the water) is

$$\mathbf{v}' = 4 \text{ mph} \times \mathbf{1}_{y'}.$$

Its motion with respect to the ground is the vector sum of the motion of the boat with respect to the water, $\mathbf{v}'$, and the motion of the water with respect to the ground, $\mathbf{V}$ $(= 3 \text{ mph} \times \mathbf{1}_x)$. According to Eq. 1-2.5, we have

$$\mathbf{v} = (3\,\mathbf{1}_x + 4\,\mathbf{1}_y) \text{ mph}.$$

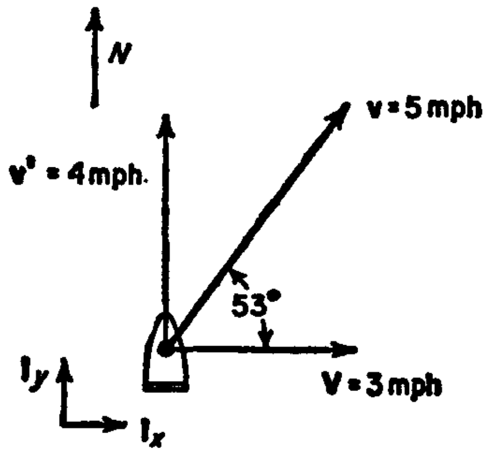


FIG. 1-2.2

The boat moves at 5 mph with respect to the ground in a direction 53° north of east.

To examine the way in which accelerations transform in two inertial frames, we may simply take time derivatives of Eq. 1-2.5, bearing in mind that $t = t'$. Thus we may differentiate $\mathbf{v}'$ with respect to t' and $\mathbf{v}$ with respect to t. Noting that $\mathbf{V}$ is constant, we find

$$\frac{d\mathbf{v}'}{dt'} = \frac{d\mathbf{v}}{dt},$$

or

$$\mathbf{a}' = \mathbf{a}. \qquad (1\text{-}2.6)$$

In this section we have been careful to emphasize that the time measured in both primed and unprimed frames has the same value, that the operation of clocks does not appear to depend on the speed with which they move. This view derives from ordinary experience which is summarized in the Galilean transformation equations. We have done this in anticipation of our study of relativity where a distinction must be made between measurements of time in two inertial frames. This distinction is one of the key problems of the special theory of relativity. While the Galilean transformation is extremely useful in the analysis of physical systems, the work of Lorentz and Einstein at the beginning of the present century showed it to be inadequate to describe the behavior of physical systems when the translational velocity

V is very large, of the order of the speed of light in vacuum (186,000 miles per second). When speeds approach the speed of light, physicists have learned that they must use a different set of transformation equations to relate motions in two inertial frames. These are called the Lorentz transformation equations. We will devote a great deal of attention to examining these equations and their consequences.

Nevertheless, although the Galilean transformation is now known to be inadequate to relate observations made on systems moving at high speeds to observations of the same systems moving slowly, it is simpler and sufficiently accurate in the limit of low velocities to be of great practical importance. Man has not yet learned to move anything but elementary particles and perhaps the nuclei of atoms at speeds approaching that of light. Thus for most practical purposes and for virtually all of engineering the Galilean transformation yields sufficiently correct results.

1-2.1 A ship heads due north at 30 mph with respect to the sea. A wind blows to the east at 40 mph with respect to the sea. Find the velocity of the wind with respect to the ship, as it might be measured by an observer on the ship. [50 mph, 37° south of east]

1-2.2 After going a mile upstream in a motor boat, a man accidentally drops an oar overboard. He proceeds upstream for 10 minutes before he misses the oar. He then turns round and retrieves the oar at the point from which he started originally. If the boat travels at constant speed with respect to the water, what is the speed of the current? Work the problem twice, in the frame fixed in the banks, and in the frame fixed in the stream. [3 mph]

1-2.3 A boy throws a stone to the east at 30 ft/sec from a northbound train traveling at 88 ft/sec on straight tracks. The stone's initial velocity is at 37° above the horizontal. The stone leaves the boy's hand at 10 ft above the ground. Find the parametric equations of motion of the stone with the point of projection as origin in a frame fixed in the ground, and in a frame fixed in the train.

§1-3 The Speed of Light In principle, measurement of the speed of light is not different from the measurement of any other speed. One measures the time required for light to travel over a measured path. The difficulty arises from its great rapidity, so that either long paths or short time intervals must be measured.

The first measurement of the speed of light was astronomical. In 1675 Olaf Roemer (1644-1710), a Danish astronomer, observed that a greater time elapsed between the eclipses of one of Jupiter's moons during the season the earth is receding from Jupiter, position E_1 in Fig. 1-3.1, than during the season the

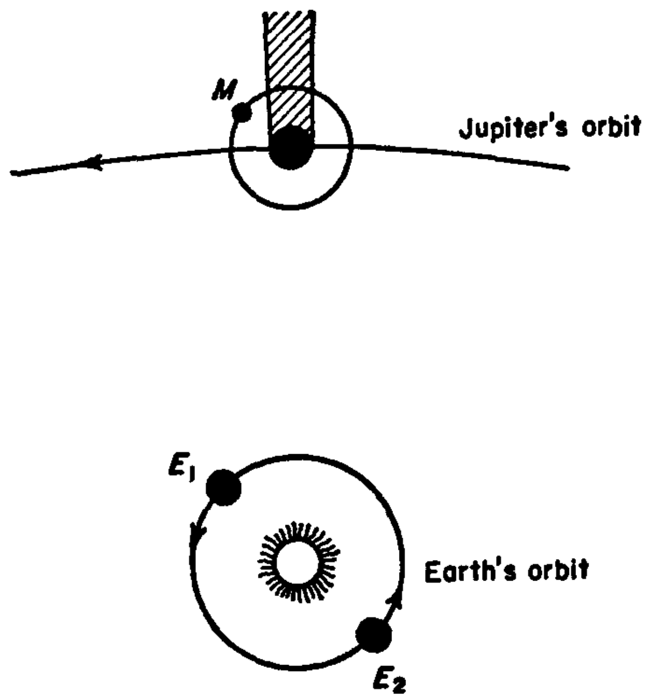


FIG. 1-3.1

earth is approaching Jupiter, position E_2 in Fig. 1-3.1. He interpreted this discrepancy as due to the finite speed of light, asserting that the light had to travel a greater distance between eclipses when the earth was receding than when the earth was approaching Jupiter. Roemer concluded that it took light 22 minutes to cross the earth's orbit, a distance of about 3×10^8 km, giving a speed of about 2.3×10^8 m/sec. Roemer's was a beautiful inference, giving a value within about 25 percent of the presently accepted value of the speed of light c, for today we take

$$c = 2.997925 \pm 0.000003 \times 10^8 \text{ m/sec.}$$

An earlier attempt to measure the speed of light was made by

Galileo, who stationed two observers with lanterns a mile apart. They were to signal each other by uncovering their lanterns according to a prearranged schedule. While such a system might have yielded a value for the speed of sound, it was unable to respond with sufficient speed to determine the speed of light. Subsequent experimenters replaced one of Galileo's observers with a mirror and replaced his crude measurement of time by the rotation of a toothed wheel. The toothed wheel was at once a shutter to let light pass and a clock, for, at the proper speed of rotation, light passing through one opening could reach the mirror and return through the next opening in the wheel. Such an arrangement was used much later by Fizeau (1819-1896) in 1849, with a base line of 8.633 km from source to mirror. In Fizeau's experiment the wheel had 720 teeth and rotated at 12.6 revolutions/sec. More modern experiments form a pulse of light (or other electromagnetic radiation) electronically, and detect it in a similar way. Today it is possible to measure time intervals nanoseconds (10^{-9}) long. But in principle the experimental measurement of the speed of light is closely related to the system used by Galileo.

Toward the end of the nineteenth century, as measurements of the speed of light became more refined, the question arose—What was the coordinate frame with respect to which the speed of light was measured? Was the speed of light which was measured its speed with respect to the earth? to the stars? to Jupiter? Was there a preferred medium for light, as air is a preferred medium for sound?

It is clear that sound travels through air, and that the speed of sound must be measured with respect to the air. Experiment shows that the velocity of sound with respect to the earth can be obtained by a Galilean transformation by adding the velocity of sound with respect to the air to the velocity of the air with respect to the earth. By analogy, physicists supposed that there was a preferred medium for light which filled infinite space. This medium was called the *ether*. The ether was thought to be perfectly transparent to light, and to offer no resistance to the passage of stars or the planets. Its function was to provide a coordinate frame to answer the question—"With respect to what is the velocity of light c?" Since the speed of the earth's orbital motion is about 30,000 m/sec, it was supposed that sensitive experiments might be devised to show that the speed of light with respect to the earth would depend on the direction the light traveled with respect to the earth's motion through the ether, in accordance to the Galilean transformation.

Consider an experimental apparatus which might be devised to study variations in the speed of light with respect to the earth, as related to the motion of the earth through space. Such an apparatus might compare the time for a light ray to pass down a rod and be reflected back, when the rod was parallel to the earth's motion, and then when the rod was perpendicular to the earth's motion.

Imagine an equipment consisting of a rod of length L provided with a mirror at one end, and with a pulsed light source, a detector, and a clock at the other end. In Fizeau's experiment the detector was the eye, and the light pulser-clock was the toothed wheel. With this idealized equipment, suitably scaled and suitably fast, we might measure the speed of light.

When the rod is *at rest* with respect to the ether, the time for the light to make a round trip is Δt_0, where

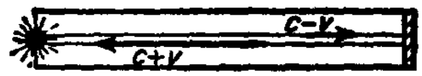
$$\Delta t_0 = 2L/c. \tag{1-3.1a}$$

When the rod is *moving* through the ether with speed v and is oriented *parallel* to its direction of motion, as in Fig. 1-3.2a, we find the speed of light with respect to the rod by the Galilean transformation, for this is the hypothesis to be tested. Then the speed of light with respect to the rod is $c - v$ when light is moving to the mirror, and $c + v$ on the return trip. The elapsed time for this trip is $\Delta t_{\parallel}$, as given by

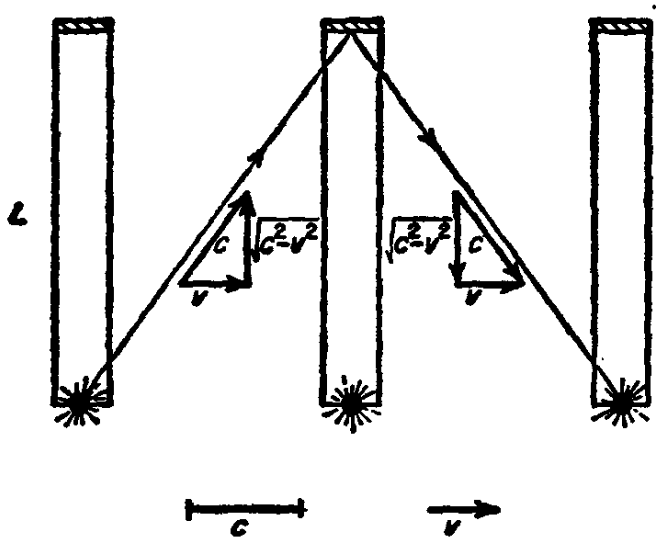
$$\Delta t_{\parallel} = L/(c - v) + L/(c + v),$$

or

$$\Delta t_{\parallel} = \frac{2L}{c}\frac{1}{(1 - v^2/c^2)}. \tag{1-3.1b}$$



(a)



(b)

FIG. 1-3.2

Suppose the rod is moving through the ether with speed v, Fig. 1-3.2b, the rod being oriented perpendicular to its direction of motion. By the Galilean transformation the speed of light with respect to the rod is $(c^2 - v^2)^{1/2}$ on both the outgoing and the return trip, so that the elapsed time is $\Delta t_\perp$, given by

$$\Delta t_\perp = \frac{2L}{c}\frac{1}{(1 - v^2/c^2)^{1/2}}. \qquad (1\text{-}3.1c)$$

If we now make use of the binomial theorem for $(1 + x)^m$,

$$(1 + x)^m = 1 + \frac{mx}{1!} + \frac{m(m-1)}{2!}x^2 + \ldots \frac{m(m-1)\ldots(m-n+2)}{(n-1)!}x^{n-1}, \qquad (1\text{-}3.2)$$

we may expand Eqs. 1-3.1b and c to the quadratic term to find

$$\Delta t_{\parallel} = \Delta t_0(1 + v^2/c^2 + \ldots) \tag{1-3.3a}$$

and

$$\Delta t_{\perp} = \Delta t_0(1 + \tfrac{1}{2}v^2/c^2 + \ldots). \tag{1-3.3b}$$

The differences in the elapsed time for the motion of light through the ether in the three cases examined are in the second order (quadratic terms) of v/c. That is, these times differ in the difference between $(v/c)^2$ and 1. For the velocities with which the earth moves in its orbit, the quantity $(v/c)^2$ is about 10^{-8}. We must measure $\Delta t_{\parallel}$ and $\Delta t_{\perp}$ to a precision of at least one part in a hundred million if two separate measurements are to be made and compared.

However, a way around the difficulty was conceived by Albert A. Michelson (1852-1931), and in 1887 Michelson and Morley performed an experiment which should have been able to detect the motion of the earth through the ether, or put in another way, should have been able to test the applicability of the Galilean transformation to the motion of light. Michelson reasoned that a direct comparison of the parallel and perpendicular transit times might be made by using light waves as their own means of measuring time. Suppose we combine the parallel and perpendicular rods, as in Fig. 1-3.3. Then a beam of light from a source is split into two by a lightly silvered mirror, half the light moving

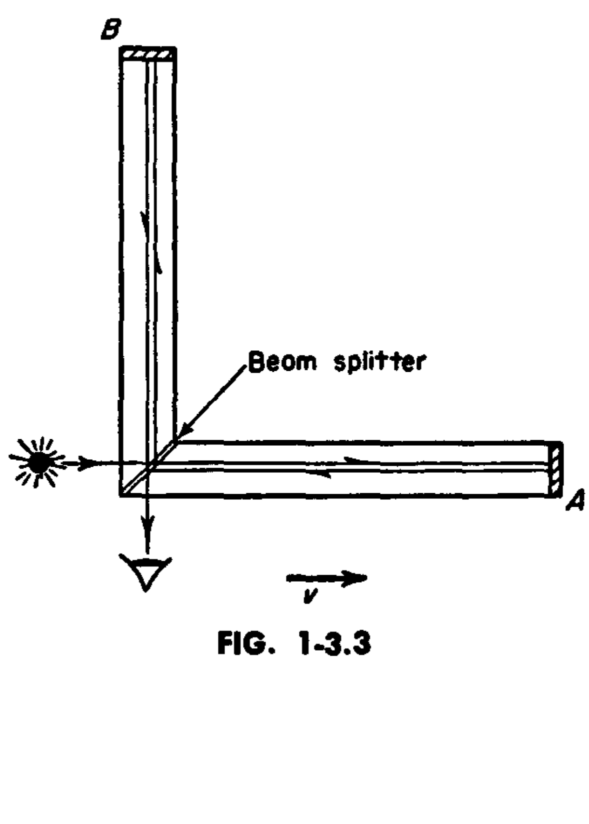


FIG. 1-3.3