5 SPECIAL THEORY OF RELATIVITY

Physics at the end of the nineteenth century looked back on a period of a great progress. The theories developed over proceeding three centuries had been very successful in explaining a wide range of natural phenomena. Newtonian mechanics beautifully explained the motion of objects on earth and in the heaven, furthermore, it formed the basis for successful treatment of fluids, wave motion and sound. Kinetic theory on the other hand, explained the behavior of gases and other materials.

Maxwell's theory of electromagnetism not only brought together and explained electric and magnetic phenomena, but it also predicted the existence of electromagnetic waves.

A few puzzles remained, but it was felt that these would soon be explained using already known principles.

But it did not turn out so simply. Instead, these puzzles were only to be solved by the introduction of two revolutionary new theories - the theory of relativity and quantum theory. Now we shall deal with the special theory of relativity, which was first proposed by Albert Einstein (1879-1955).

5.1 Galilean and Newtonian Relativity

Einstein's special theory of relativity deals with how we observe events, particularly how objects and events are observed from different frames of reference. A reference frame is a set of coordinate axis fixed to some body.

We will deal with so-called **inertial reference frames**. An inertial reference frame is one in which Newton s first law, the law of inertia is valid. That is if an object experiences no net force due to other bodies, the object either remains at rest or in motion with constant velocity in a straight line. Rotating or otherwise accelerating frames of reference are non-inertial frames.

So a reference frame that moves with constant velocity with respect to another inertial frame is itself also an inertial frame.

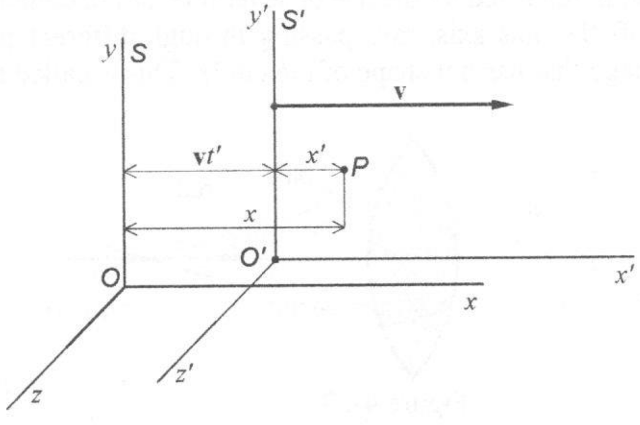

Figure 5-1

Now we examine in detail the mathematics of relating quantities in one inertial reference frame to the equivalent quantities in another. In particular we will see how positions and velocities transform from one reference frame to the other.

Let us consider two reference frames S and S′, see Fig.5-1. The x and x' axes overlap one another. We also assume that frame S′ moves to the right (in the x direction) at speed v

with respect to reference frame S. For the sake of simplicity let us assume that the origins O and O′ are superimposed at time $t = 0$.
Now consider an event that occurs at some point P represented by the coordinates x', y', z' in reference frame S′ at the time t'. What will be the coordinates of point P in S reference frame? Since S and S′ overlap precisely initially, after a time t', the frame S′ will have moved a distance (vt'). Therefore, at time t′ we can write

$$x = x' + vt'. \qquad (5\text{-}1)$$

The y and z coordinates, on the other hand, are not altered by motion along the x-axis, thus

$$y = y' \quad \text{and} \quad z = z'. \qquad (5\text{-}2)$$

Finally since time is assumed to be absolute in Newtonian physics, clocks in the two frames will agree with each other, so $t = t'$.
We summarize these in the following **Galilean transformation equations**:

$$x = x' + vt' \qquad y = y' \qquad z = z' \qquad t = t'. \qquad (5\text{-}3)$$

These equations give the coordinates of an event in the S frame when those in the S′ frame are given. If those in the S frame are known, then we can obtain very easily **inverse transformation**:

$$x' = x - vt \qquad y' = y \qquad z' = z \qquad t' = t. \qquad (5\text{-}4)$$

Now suppose that a point P in Fig.5-1 represents a particle that is moving. Let the components of its velocity vector in S' be u_x', u_y', u_z'. Now

$$u_x' = \frac{dx'}{dt} \qquad u_y' = \frac{dy'}{dt} \qquad u_z' = \frac{dz'}{dt}. \qquad (5\text{-}5)$$

The velocity of P as seen from S reference frame will have components u_x, u_y and u_z. Let us show how these are related to the velocity components in S' by differentiating Galilean transformation equations:

$$u_x = \frac{dx}{dt} = \frac{d}{dt}(x' + vt) = \frac{dx'}{dt} + v = u_x' + v$$

$$u_y = \frac{dy}{dt} = \frac{dy'}{dt} = u_y' \qquad (5\text{-}6)$$

$$u_z = u_z'.$$

These equations are known as **Galilean velocity transformation**.
Both Galileo and Newton were deeply aware of what we now call the **relativity principle** - that is **the basic laws of physics are the same in all inertial reference frames**. You may have recognized its validity in everyday life: for example, objects move in the same way in a smoothly moving train (constant velocity) as they do on earth. When you play Ping-Pong while riding in such a train the ball moves just as it does on the earth.
Newton's physics, which we used until now, involves certain unprovable assumptions that make sense from everyday experience. It is assumed that the lengths of objects are the same in one reference frame as in another, and that time passes at the same rate in different reference frames. In classical mechanics space and time we consider being absolute, or electric charge is assumed to be unchanged by a change in inertial reference frame.

The fact that the laws of mechanics are the same in all inertial reference frames implies that: **no one inertial reference frame is special in any sense**. We express this important conclusion by saying that **all inertial reference frames are equivalent** for the description of mechanical phenomena. No one inertial reference frame is any better than another is. When you travel smoothly at constant velocity in a train, it is just as valid as to say that you are at rest and the earth is moving in reverse direction. There is no experiment one can do to say which frame is „really" at rest and which one is moving.

The situation changed somewhat when Maxwell presented his theory of electromagnetism. He showed that light could be considered as electromagnetic waves. Maxwell's equations predicted that the velocity of light would be 3×10^8 m/s and this is just what was measured within experimental error. The question then arose: in what reference frame does light have precisely the velocity predicted by Maxwell's theory. For it was assumed in agreement with Galilean velocity transformation equations, that light would have a different speed in different frames of reference.

For example if observers were traveling on a rocket at a speed of 10^8 m/s toward a source of light, they will measure the speed of light reaching them to be $1\times10^8 + 3\times10^8 = 4\times10^8$ m/s.

But Maxwell's equations have no provision for relative velocity. They merely predicted the speed of light to be $c = 3\times10^8$ m/s. This seemed to imply that there must be some special reference frame where c would have this value.

We know that light can be considered as electromagnetic waves. But we also know that waves travel on water and along ropes and strings, sound waves travel in air and in other materials. Since nineteenth century physicists viewed the material world in terms of the laws of mechanics, it was natural for them to assume that light must travel in some medium. They called this transparent medium the ether and assumed it permeated all space. And it was therefore assumed that the velocity of light given by Maxwell's equations must be with respect to this ether.

Scientists soon set out to determine the speed of the earth relative to this absolute reference frame whatever it might be. A.A. Michelson and E.W. Morley performed the most famous experiment.

5.2 Michelson - Morley Experiment, The Postulates

This experiment was designed to measure the speed of ether, the medium in which light was assumed to travel with respect to the earth.

One of the possibilities that nineteenth centuries considered was that the ether is fixed to the sun, for even Newton had taken the sun as the center of the universe. If this were the case (there was no guarantee of course) the earth speed of about 3×10^4 m/s in its orbit around the sun would produce a change 1 part in 10^4 in the speed of light.

Michelson and Morley assumed that the motion of the earth with respect to the stationary ether will produce a time lag between two beams traveling parallel and perpendicular to the motion of the earth, see Fig. 5-2. They were able to measure this time lag using the principle of interference. The time lag depends on the velocity of the earth with respect to the ether.

We imagine that earth is at rest. So the ether wind moves to the left with the speed v. The light from the source falls on the semitransparent mirror and is divided into two beams:

- 1st beam, traveling perpendicular to the ether wind,
- 2nd beam, traveling parallel to the ether wind.

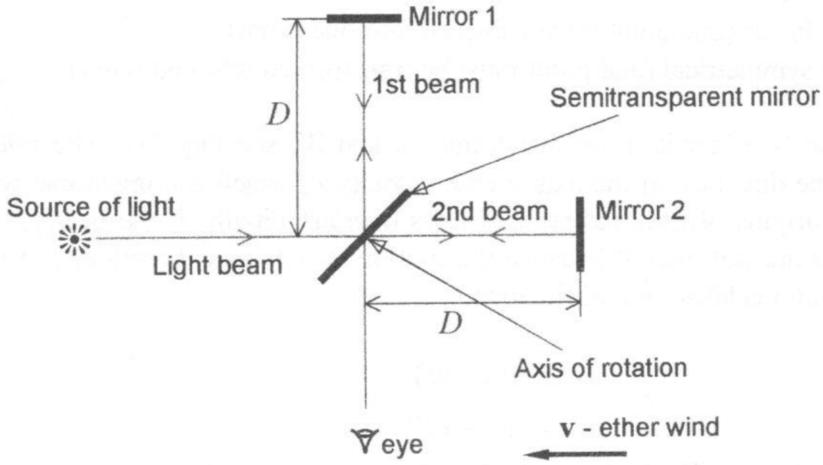

Figure 5-2

After reflection of these two beams on mirrors 1 and 2 the beams interfere so that the observer can see the interference pattern. On the base of the interference pattern shift, when the interferometer is rotated, it is possible to determine the time lag between the beams. But Michelson and Morley did not find a significant interference pattern shift.
The man who explained this great puzzle of physics was Albert Einstein. In his paper from 1905 he proposed doing away completely with the idea of ether and the accompanying assumption of an absolute reference frame at rest. This assumption was embodied in two postulates. First postulate is called **the relativity principle**:

The laws of physics have the same form in all inertial reference frames.

Second postulate deals with **the constancy of the speed of light**:

Light propagates through empty space with a definite speed *c* independent of the speed of the source or observer.

This second postulate is a direct result of the null result of the Michelson experiment. Thus a person traveling toward or away from a source of light will measure the same speed for that light as someone at rest with respect to the source (in vacuum). This conflicts with our everyday notions, for we would expect to have to add the velocity of the observer.
Thus we can say that the second principle of special theory of relativity is in contradiction with Galilean velocity transformation. This transformation is valid only when velocities involved are much less than the speed of light.
Clearly a new set of transformation equations is needed to deal with relativistic velocities. This new set of equations is called the Lorentz transformation equations.

5.3 Lorentz Transformation

We will now derive new transformation which:
- will be in agreement with both principles of special theory of relativity,

- for the small velocities must change into Galilean transformation,
- must be linear (one point is transformed into one point),
- must be symmetrical (end point must be transformed into end point).

We imagine two inertial reference frames S and S′, see Fig. 5-1. The reference frame S′ moves in the direction of the axis x with velocity v, which is comparable with the speed of light. The origins of both reference frames overlap initially at the time $t = 0$. The y and z coordinates are not altered because the motion is only in x direction. We will assume that transformation is linear and of the form

$$x' = \gamma(x - vt), \tag{5-7}$$

$$x = \gamma(x' + vt'). \tag{5-8}$$

The constant γ is to be determined. We will not assume a form for t, but we will derive it. Now if a light pulse leaves the common origin O and O′ at time $t = t' = 0$ after a time t it will have traveled along the axis x a distance

$$x = ct, \tag{5-9}$$

or in printed frame $$x' = ct'. \tag{5-10}$$

Substituting Eqs.5-7 and 5-8 into Eqs.5- 9 and 5-10 we have

$$x = ct = \gamma(ct' + vt') = \gamma(c + v)t', \tag{5-11}$$

$$x' = ct' = \gamma(ct - vt) = \gamma(c - v)t. \tag{5-12}$$

We substitute t' from Eq.5-12 into Eq.5-11, or

$$x = ct = \gamma(c+v)\,\frac{\gamma(c-v)t}{c} = \gamma^2\,\frac{(c^2 - v^2)t}{c}.$$

Solving for γ we have

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{5-13}$$

Now that we have found γ, we need only find the relation between t and t'. To do so we combine Eq.5-7 with Eq.5-8, or

$$x' = \gamma(x - vt) = \gamma\left[\gamma(x' + vt') - vt\right].$$

Solving for t we obtain

$$t = \gamma\left[t' - \frac{x'}{v}\left(\frac{1}{\gamma^2} - 1\right)\right].$$

Taking into account Eq.5-13 we obtain for time transformation

$$t = \frac{t' + \frac{x'v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{5-14}$$

In summary we can write **Lorentz transformation**

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad y' = y, \qquad z' = z, \qquad t' = \frac{t - \frac{xv}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{5-15}$$

After a little rearrangement we can also write for **inverse transformation**

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad y = y', \qquad z = z', \qquad t = \frac{t' + \frac{x'v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{5-16}$$

Notice that not only is the x equation modified as compared to the Galilean transformation, but so is the t equation, indeed we see directly in this last equation how the space and time coordinates mix.
We can see an important consequence of the special theory of relativity namely that the time can no longer be regarded as absolute quantity. From the transformation equations we see that time depends also on the observer's position. Now let us have a closer look on another consequences of special theory of relativity.

5.4 Length Contraction

As we have already mentioned the time has lost its special, absolute position and now we know that it depends on the reference frame in which certain events take place. Similarly not only time intervals between two events are different in different reference frames. The lengths and distances are different as well in different reference frames.
Imagine we have two different reference frames S and S′, see Fig. 5-3.
Let us consider a rod, which is at rest along the axis x' in the reference frame S′. This reference frame moves with respect to the reference frame S with speed v along x axis. The length of the rod in the reference frame S′ is denoted as l_0.
Thus we have

$$l_0 = x_2' - x_1'. \tag{5-17}$$

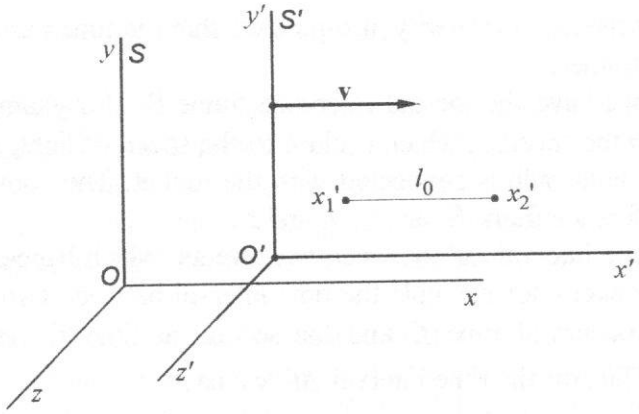

Figure 5-3

In Eq. 5-17 x_2' and x_1' are coordinates of its endpoints at the same instant of time. At any instant of time the coordinates x_2 and x_1 of the endpoints of the rod in the reference frame S could be determined from Lorentz transformation. Thus we can express the length l of the rod measured in the unprimed reference frame S as

$$l = x_2 - x_1, \tag{5-18}$$

where x_2 and x_1 are coordinates of its endpoints at the same instant of time. To find the length of the rod we substitute Lorentz transformation (see Eq.5-15) into Eq.5-17. Thus we obtain

$$l_0 = \frac{x_2 - vt}{\sqrt{1-\frac{v^2}{c^2}}} - \frac{x_1 - vt}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{l}{\sqrt{1-\frac{v^2}{c^2}}},$$

or

$$l = l_0\sqrt{1-\frac{v^2}{c^2}}. \tag{5-19}$$

This equation gives the length that will be measured when an object travels past an observer at speed v. It is important to note that the length contraction occurs only along the direction of motion. The length contraction is a general result of the special theory of relativity and applies to lengths of objects as well as to distances. This result can be stated most simply in words as:

The length of an object is measured to be shorter when it is moving than when it is at rest.

The length l_0 is called the **proper length**. It is the length of the object as measured by an observer at rest with respect to it.

As far as we have the speed of the reference frame S′ in the expression for length contraction squared it does not matter which of the systems is denoted as S and which of them is denoted as S′. What is important that l_0 – the proper length - is the length of the object in the reference frame in which the body is at rest.

5.5 Time Dilatation

The Einstein's special theory of relativity also predicts that the time passes at different rates in different reference frames.

Let us suppose that we have the inertial reference frame S′ (for example connected with a rocket) moving with the speed v, which is close to the speed of light, see Fig. 5-4. Let us also suppose that the light bulb is connected with the rocket. The coordinates of the light bulb in the moving reference frame S′ are x', y' and z'.

Let us measure the time interval $\Delta t'$ between two events, which happen at the same place on the board of the rocket - for example the time interval between two flashes of the light bulb. The first flash occurs at time t_1' and the second at time t_2' (measured in moving reference frame S′). Thus for the time interval $\Delta t'$ we have

$$\Delta t' = t_2' - t_1'. \tag{5-20}$$

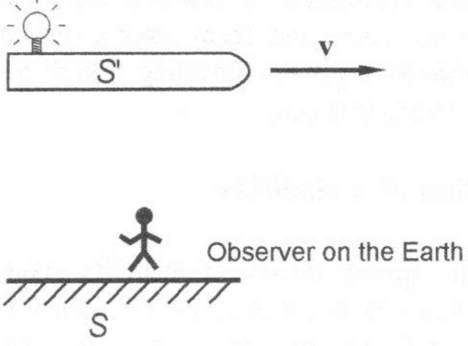

Figure 5-4

Let us now find the time interval Δt between these two flashes measured by an observer on the earth – in the reference frame S. To do so we use Lorentz transformation for the instant of time t_2' and t_1', so that the time interval in the reference frame S will be

$$\Delta t = t_2 - t_1 . \tag{5-21}$$

Note that the events – flashes of light bulb – take place at one point x'. Substituting Lorentz transformation for time (see Eq.5-16) into Eq.5-21 we have

$$\Delta t = \frac{t_2' + \frac{x'v}{c^2} - t_1' - \frac{x'v}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{t_2' - t_1'}{\sqrt{1-\frac{v^2}{c^2}}},$$

or

$$\Delta t' = \Delta t\sqrt{1-\frac{v^2}{c^2}} . \tag{5-22}$$

Since $\sqrt{1-\frac{v^2}{c^2}}$ is always less than one, we see that $\Delta t > \Delta t'$. That is the time between two events which take place at the same point is greater for the observer on the earth than for the traveling observer. This is a general result of special theory of relativity, which is known as time dilatation. Stated simply the time dilatation effect says that:

Moving clocks are measured to run slowly.

Time is actually measured to pass slowly in any moving reference frame as compared to your own. The concept of time dilatation may be hard to accept for it violates our common sense. As we can see from Eq.5-22 the time dilatation effect is negligible unless v is reasonable close to the speed of light. The speeds we experience in everyday life are very much smaller than speed of light, so it is little wonder that we do not ordinarily notice the time dilatation.

Experiments have been done to test the time dilatation effect and to confirm Einstein's predictions. In 1971, extremely precise atomic clocks were flown around the earth in a jet plane. Since the speed of the plane (about 1000 km/h) is much less than speed of light, the clocks had to be accurate to 10^{-9} s in order to detect the time dilatation. The result of experiment confirmed the prediction.

Time dilatation has been also confirmed by studying the lifetime of elementary particles. Many of these particles are not stable and decay after a certain time into simpler particles. Careful experiments show that for a particle traveling at high speed its lifetime increases just as predicted by the time dilatation formula.

5.6 Relativistic Addition of Velocities

The second principle of the special theory of relativity states that the speed of light is the same is all inertial reference frames. Let us have a look if this principle is in agreement with the Galilean velocity transformation. Suppose the following example, see Fig. 5-5.

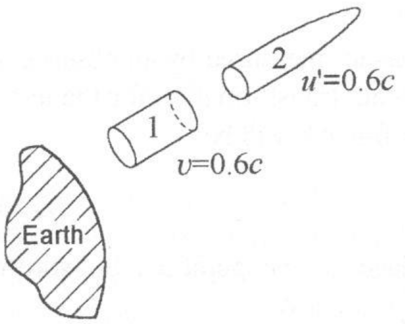

Figure 5-5

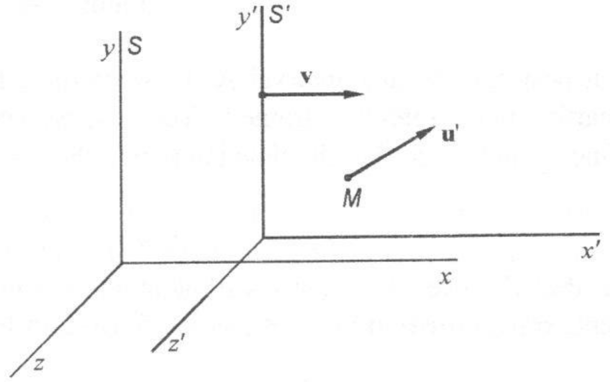

Figure 5-6

The first rocket moves with the speed $\upsilon = 0.6\ c$ with respect to the earth. Rocket 2 is fired from rocket 1 with speed $u' = 0.6\ c$ (with respect to rocket 1). Following Galilean velocity transformation the speed of the rocket 2 with respect to the earth will be 1.2 c and this is in contradiction with the second principle of the special theory of relativity.
Relativistically correct velocity transformation equations can be obtained from Lorentz transformation. Let us suppose that we have two reference frames S and S′, see Fig. 5-6. The reference frame S′ moves with respect to the reference frame S with speed υ close to the speed of light. Let the particle M moves with respect to the reference frame S′ with velocity u'. We would like to determine the velocity u of the particle as measured in the frame S.
Measured in the S frame the velocity vector of the particle has components

$$u_x = \frac{dx}{dt}, \qquad u_y = \frac{dy}{dt}, \qquad u_x = \frac{dz}{dt}. \qquad (5\text{-}23)$$

The velocity vector, as measured in the S′ frame has components

$$u'_x = \frac{dx'}{dt}, \qquad u'_y = \frac{dy'}{dt}, \qquad u'_z = \frac{dz'}{dt}. \qquad (5\text{-}24)$$

To establish the required relationships, we take the differentials of the Lorentz transformation (see Eq.5-16) remembering that υ is constant. This gives

$$dx = \frac{dx' + v dt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad dy = dy', \qquad dz = dz', \qquad dt = \frac{dt' + \frac{v}{c^2} dx'}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{5-25}$$

Thus we obtain

$$u_x = \frac{dx}{dt} = \frac{\dfrac{dx' + v dt'}{\sqrt{1 - \frac{v^2}{c^2}}}}{\dfrac{dt' + \frac{v}{c^2} dx'}{\sqrt{1 - \frac{v^2}{c^2}}}} = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}}, \qquad u_y = \frac{dy}{dt} = \frac{\sqrt{1 - \frac{v^2}{c^2}}\, u'_y}{1 + \frac{u'_x v}{c^2}},$$

$$u_z = \frac{dz}{dt} = \frac{\sqrt{1 - \frac{v^2}{c^2}}\, u'_z}{1 + \frac{u'_x v}{c^2}}. \tag{5-26}$$

These equations constitute the relativistic velocity transformation, which is in agreement with both principles of special theory of relativity. Note that as v/c approaches zero equations 5-26 approach those derived from the Galilean transformation – see Eq.5-6.

5.7 Relativistic Mass, Momentum and Energy

When a steady net force is applied to an object, the object increases its speed. Following second Newton's law $\mathbf{F} = m\mathbf{a}$ the speed of the object can increase indefinitely, even to the values exceeding the speed of light. This is, however, in contradiction with second principle of the special theory of relativity.
Relativistically valid Newton's second law is therefore stated in the following form

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}), \tag{5-27}$$

where mass of the object is no longer constant but it is a function of its speed. We assume that the momentum of the object is

$$\mathbf{p} = m\mathbf{v}, \tag{5-28}$$

which is just like the classical formula for momentum, except that m is a function of a speed of the object. From the condition of validity of conservation of momentum law in the relativistic domain it can be shown that the dependence of the mass on the speed is

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{5-29}$$

where m_0 is called the **rest mass** of the object – the mass it has as measured in a reference frame in which it is at rest and m is the mass it will be measured to have when it moves at speed v.

The relativistic mass increase has been tested many times on elementary particles, and the mass has been found to increase in accord with the previous formula. Taking into account dependence of mass on speed we can express relativistic momentum as

$$\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{5-30}$$

To find a mathematical relationship between mass and energy we assume that the work-energy theorem is still valid in relativity and we take the motion to be along the x-axis. Work-energy theorem states that the net work done on the object is equal to its change in kinetic energy.

The work done to increase the object speed from zero to v is

$$W = \int F dx = \int \frac{dp}{dt} dx = \int \frac{dp}{dt} v dt = \int v dp. \tag{5-31}$$

Since $d(pv) = pdv + vdp$ we can write

$$vdp = d(pv) - pdv. \tag{5-32}$$

Substituting Eq.5-32 into Eq.5-31 we have

$$W = \int vdp = \int d(pv) - \int pdv.$$

Since integration is the exact inverse of differentiation the term $\int d(pv)$ becomes

$$\int_0^v d(pv) = pv\Big|_0^v = mv^2\Big|_0^v = mv^2,$$

where m is a function of v. Therefore we have

$$W = mv^2 - \int pdv = mv^2 - \int mvdv = mv^2 - \int_0^v \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}} dv. \tag{5-33}$$

The second term is easily integrated. We denote $\mathrm{I} = \int_0^v \frac{m_0 v}{\sqrt{1-\frac{v^2}{c^2}}} dv$. Using the substitution

$u = \sqrt{1-\frac{v^2}{c^2}}$, we obtain $\frac{vdv}{\sqrt{1-\frac{v^2}{c^2}}} = -c^2 du$. Thus we can write the integral I as

$$\mathrm{I} = m_0\left[-c^2 \int du\right] = -m_0 c^2\left[\sqrt{1-\frac{v^2}{c^2}} - 1\right]. \tag{5-34}$$

Substituting Eq.5-34 into Eq.5-33 we have after little rearrangement

$$W = m_0 c^2 \left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1 \right). \tag{5-35}$$

By the work-energy theorem the work done must be equal the final kinetic energy since the object started from rest. Therefore for kinetic energy we have

$$E_K = mc^2 - m_0c^2 = m_0c^2\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} - 1\right). \tag{5-36}$$

Note that for speed $v << c$ the expression for E_K must go into a classical formula $\frac{1}{2}mv^2$.

To show this we use a binomial expansion for term $\left(1-\frac{v^2}{c^2}\right)^{-\frac{1}{2}}$. Substituting this expansion into Eq.5-36 se obtain

$$E_K = m_0c^2\left(1+\frac{1}{2}\frac{v^2}{c^2}+\ldots-1\right) \approx \frac{1}{2}m_0v^2. \tag{5-37}$$

The equation for $E_K = mc^2 - m_0c^2$ requires some interpretation. First of all what does the second term m_0c^2 mean? Consistent with the idea that mass is a form of energy, Einstein called m_0c^2 the **rest energy of the object**. We can rearrange the equation for kinetic energy to get

$$mc^2 = m_0c^2 + E_K. \tag{5-38}$$

We call mc^2 the **total energy** E of the object

$$E = mc^2. \tag{5-39}$$

Thus we see that the total energy equals the rest energy plus the kinetic energy, which is Einstein's famous formula. For a particle at rest its total energy is

$$E_0 = m_0c^2, \tag{5-40}$$

which we have called its rest energy. This formula relates the concepts of energy and mass. Thus the mass ought to be convertible to energy and vice versa. That is if mass is just one form of energy then it should be converted to other forms. This is the best known result of special theory of relativity called **mass-energy equivalence**. The interconversion of mass and energy can be detected in nuclear and elementary particle physics. For example the neutral pion π^0 of rest mass 2.4×10^{-28} kg is observed to decay in pure electromagnetic radiation (photons) in such a way that π^0 completely disappears in the process. The amount of electromagnetic energy produced is found to be exactly equal to that predicted by Einstein's formula.

The reverse process is always observed - for example electromagnetic radiation under certain condition can be converted into material particles, such as electrons. On a larger scale the energy produced in nuclear power plants is a result of the loss in mass of the uranium fuel as it undergoes the process called fission. Even the radiant energy we receive from the sun is an instance of $E = mc^2$; the sun's mass is continuously decreasing as it radiates energy outward.

A useful relation between the total energy E of a particle and its momentum can be also derived. Since $E = mc^2$ and $p = mv$, where $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$, we can easily obtain

$$E = \sqrt{p^2c^2 + m_0c^4}. \tag{5-41}$$

5.8 Spacetime and Four-Vectors

Einstein regarded events as the basic data of physics. He found that every inertial observer has his own privately valid time and correspondingly his own "instantaneous three space" consisting of all events (x,y,z,t) with fixed time coordinate. But it was the mathematician Minkowski (in 1908) who proposed to think of the totality of events in the world as the "points" of an absolute four-dimensional manifold, called spacetime. Minkowski's diagram, which represents spacetime, with two dimensions suppressed, is shown in Fig. 5-7. The "moments" in this graph have equation t = constant and correspond to horizontal line $a - b$, while the "history" of each fixed point on the spatial x axis corresponds to a vertical line $c - d$. Different inertial observers draw different sections through spacetime.

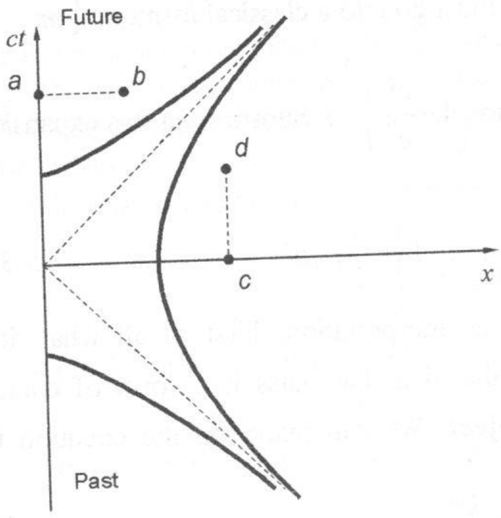

Figure 5-7

But in the relativistic case spacetime is very much more than a conventional scheme for drawing graphs. It is a four-dimensional metric space. Analogously to distance we can define for any of its two points (that is for two events A and B) the interval as

$$\Delta s_{AB} = \left[c^2(t_B - t_A)^2 - (x_B - x_A)^2 - (y_B - y_A)^2 - (z_B - z_A)^2\right]^{\frac{1}{2}}, \tag{5-42}$$

or

$$\Delta s_{AB} = \left[c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2\right]^{\frac{1}{2}}. \tag{5-43}$$

Denoting the spatial separation as

$$\Delta l = \left[\Delta x^2 + \Delta y^2 + \Delta z^2\right]^{\frac{1}{2}},$$

we can write

$$\Delta s_{AB} = \left[c^2\Delta t^2 - \Delta l^2\right]^{\frac{1}{2}}. \tag{5-44}$$

We can also express an interval in the differential form as

$$d s = \left[c^2 dt^2 - dl^2\right]^{\frac{1}{2}}. \tag{5-45}$$

The interval is absolute that is, it has the same value when evaluated by any inertial observer. We may discuss here briefly the physical significance of Δs_{AB} or better Δs^2_{AB}.

Evidently $\Delta s^2_{AB} = 0$ holds for two given events A and B if and only if these events are connectable by a light signal.

When $\Delta s^2_{AB} > 0$ then in any inertial reference frame (see Eq.5-44), $c^2 = \frac{\Delta l^2}{\Delta t^2}$, and thus an observer moving with uniform velocity less than speed of light can be sent from one of events to the other.

Similarly, if $\Delta s^2_{AB} < 0$, then $c^2 < \frac{\Delta l^2}{\Delta t^2}$ and the spatial separation Δl between events A and B is greater than the path that could be traveled by light is time Δt. Therefore there could not be any connection between events A and B.

Minkowski demonstrated that the existence of the "metric" of spacetime has a significant mathematical consequence: it leads to the four-vector calculus beautifully adapted to the needs of special theory of relativity. Relativity considers the four components, which make up a four-vector to make up a single physical quantity.

Thus we introduce instead of spatial coordinates x,y,z and time coordinate t four components x_1, x_2, x_3 and x_4 of the four-vector in the following way

$$x_1 = x \qquad x_2 = y \qquad x_3 = z \qquad x_4 = jct, \tag{5-46}$$

where j is the imaginary unit.

As an example we can express in the Minkowski spacetime an interval using Eq.5-45. The three spatial components form an ordinary vector in the three-dimensional space. Thus we can write for the spatial separation of two events

$$dl^2 = dx^2 + dy^2 + dz^2 = dx_1^2 + dx_2^2 + dx_3^2 . \tag{5-47}$$

For the temporal separation we obtain

$$dx_4^2 = -c^2 dt^2 . \tag{5-48}$$

Substituting Eqs.5-47 and 5-48 into Eq.5-45 we have for interval

$$d\,s^2 = -dx_4^2 - dx_1^2 - dx_2^2 - dx_3^2 ,$$

or

$$-d\,s^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2 = \sum_{\alpha=1}^{4} dx_\alpha dx_\alpha . \tag{5-49}$$

In analogy with distance in three-dimensional space we can consider the interval as "distance" between two events in the Minkowski spacetime. An interesting property of the interval is that it is not changed when transformed according to Lorentz transformation. Such a quantity is called a Lorentz-invariant

$$d\,s = d\,s' . \tag{5-50}$$

We can also rewrite Lorentz transformation into Minkowski spacetime. To do this we can write

$$x_1 = x, \qquad x_2 = y, \qquad x_3 = z, \qquad x_4 = jct,$$

and

$$x_1' = x', \qquad x_2' = y', \qquad x_3' = z', \qquad x_4' = jct'.$$

Substituting these equations into Eq.5-16 we obtain after little rearrangement

$$x_1 = \frac{x_1' - \frac{jv}{c} x_4'}{\sqrt{1 - \frac{v^2}{c^2}}}, \qquad x_2 = x_2', \qquad x_3 = x_3', \qquad x_4 = \frac{x_4' + \frac{jv}{c} x_1'}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{5-51}$$

These equations are Lorentz transformation for the components of any four-vector. It can be also shown that Lorentz transformation may be thought of as a rotation in the four-dimensional spacetime, when the fourth dimension is jct. To do so we rewrite Eq.5-51 into the following form

$$x_1 = \gamma x_1' - j\gamma\beta x_4', \qquad x_2 = x_2', \qquad x_3 = x_3', \qquad x_4 = \gamma x_4' + j\gamma\beta x_1', \tag{5-52}$$

where $\beta = \frac{v}{c}$ and γ is given by Eq.5-13. Due to the fact that $\gamma^2 = (j\beta\gamma)^2 = 1$ we can denote

$$\gamma = \cos\varphi \qquad j\beta\gamma = \sin\varphi .$$

Thus we can rewrite Eq.5-52 as

$$x_1 = x_1'\cos\varphi - x_4'\sin\varphi \qquad x_2 = x_2', \qquad x_3 = x_3', \qquad x_4 = x_4'\cos\varphi + x_1'\sin\varphi . \tag{5-53}$$

Figure 5-8 shows a point A in the (X_1, X_4) coordinate system. In this coordinate system the point A is said to have coordinates x_1 and x_4. The same point has the coordinates x_1' and x_4' in the coordinate system (X_1', X_4') rotated with respect to the first by the angle φ.

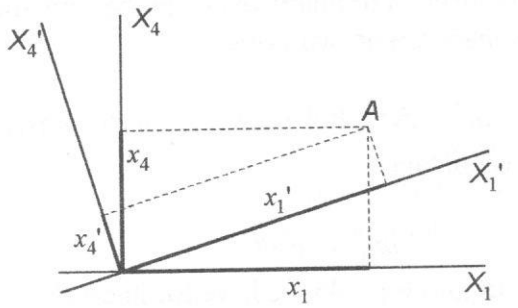

Figure 5-8

From this figure we can obtain the transformation equations

$$x_1 = x_1'\cos\varphi - x_4'\sin\varphi \qquad x_4 = x_4'\cos\varphi + x_1'\sin\varphi . \tag{5-54}$$

On the base of the formal analogy of the first and last equations of Eq.5-53 and Eq.5-54 we can therefore consider Lorentz transformation as a rotation in the four-dimensional space about axis $x_2 = x_2'$ and $x_3 = x_3'$.

To introduce four-velocity and other four-vectors we need another important invariant – so called proper time. Let us imagine we have a clock moving with respect to the inertial reference frame S. At the time dt the clock travels the path $dl = \sqrt{dx^2 + dy^2 + dz^2}$ (measured in S reference frame). We are interested in time $d\tau$ measured by the clock in the reference frame S' which is moving together with the clock. In this case $dx' = dy' = dz' = 0$. Taking into account that the interval is invariant to the Lorentz transformation, see Eq.5-50, we can write

$$d\text{s} = \sqrt{c^2dt^2 - dl^2} = cd\tau , \tag{5-55}$$

or

$$d\tau = \frac{d\text{s}}{c} = \frac{1}{c}\sqrt{c^2dt^2 - dl^2} . \tag{5-56}$$

Following the second postulate of the special theory of relativity the speed of light is also Lorentz invariant. Thus $\frac{ds}{c}$ is invariant and therefore $d\tau$ is also Lorentz invariant. The time $d\tau$ is called the proper time and it is the time measured by clock moving together with the object. We shall not be surprised, therefore, to find $d\tau$ appearing in many relativistic formulae where in a classical analogy there is a *dt*.
We can also rewrite Eq.5-56 into the following form

$$d\tau = dt\sqrt{1-\frac{1}{c^2}\left[\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2\right]}.$$

Supposing that the clock moves along *x* axis only we obtain

$$c\tau = dt\sqrt{1-\frac{\upsilon^2}{c^2}}. \qquad (5\text{-}57)$$

We have therefore arrived in different way to the expression for time dilatation – see Eq. 5-22.
Newton's mechanics is not Lorentz-invariant. It was therefore necessary to construct a new mechanics, called "relativistic" for objects moving with high speeds. In this mechanics we work with four-vectors. Therefore as an example we introduce some basic four-vectors which allow us to formulate Lorentz-invariant Newton's law.
The prototype of a four-vector is the displacement four-vector between two events. The components dx_α, α=1,2,3 and 4 of this four-vector could be obtained by differentiating Eqs.5-46.
The components of four-velocity are defined as

$$u_\alpha = \frac{dx_\alpha}{d\tau}, \qquad \alpha=1,2,3 \text{ and } 4. \qquad (5\text{-}58)$$

where $d\tau$ is the proper time. Substituting for dx_α and for $d\tau$ we have for components of four-velocity

$$u_\alpha = \frac{\upsilon_1}{\sqrt{1-\beta^2}};\frac{\upsilon_2}{\sqrt{1-\beta^2}};\frac{\upsilon_3}{\sqrt{1-\beta^2}};\frac{jc}{\sqrt{1-\beta^2}}. \qquad (5\text{-}59)$$

For the norm, or magnitude of four-velocity we easily obtain

$$u_\alpha^2 = -c^2. \qquad (5\text{-}60)$$

As far as the speed of light is Lorentz-invariant the norm of the four-velocity is also Lorentz-invariant. We can also recognize in the first three components of a four-velocity the components of a familiar three-vector velocity.
Now let us suppose that associated with each particle there is a scalar – the rest mass m_0. It turns out to be identical with the mass the particle manifests in slow motion experiments. We then define for each particle analogously to Newton's momentum a four-momentum

$$p_\alpha = m_0 u_\alpha, \qquad \alpha=1,2,3 \text{ and } 4. \qquad (5\text{-}61)$$

Using Eq.5-59 we obtain for components of the four-momentum

$$p_\alpha = \frac{m_0 v_1}{\sqrt{1-\beta^2}}; \frac{m_0 v_2}{\sqrt{1-\beta^2}}; \frac{m_0 v_3}{\sqrt{1-\beta^2}}; \frac{j m_0 c}{\sqrt{1-\beta^2}}. \tag{5-62}$$

We can show, using Eqs.5-29 and 5-39 that the fourth component is proportional to the total energy of the particle

$$p_4 = \frac{j m_0 c}{\sqrt{1-\beta^2}} = \frac{j}{c}\frac{m_0 c^2}{\sqrt{1-\beta^2}} = \frac{j}{c}E. \tag{5-63}$$

Thus we can write for components of the four-momentum

$$p_\alpha = \left(p_1, p_2, p_3, \frac{j}{c}E\right). \tag{5-64}$$

Taking into account Eq.5-60 we can express the norm of four-momentum

$$p_\alpha^2 = m_0^2 u_\alpha^2 = -m_0^2 c^2. \tag{5-65}$$

However we can express the norm of four-momentum also from Eq.5-64. Thus we obtain

$$p_\alpha^2 = p^2 - \frac{E^2}{c^2}, \tag{5-66}$$

where p is an ordinary momentum. From comparison of Eqs.5-65 and 5-66 we can obtain well - known formula relating energy and momentum of the particle

$$E = \sqrt{p^2c^2 + m_0 c^4}.$$

We can now define the four-force on a particle having four-momentum with components p_α by the relation

$$F_\alpha = \frac{dp_\alpha}{d\tau}, \qquad \alpha = 1,2,3 \text{ and } 4. \tag{5-67}$$

This equation is the Lorentz-invariant Newton's law in the four-dimensional space.
The components of all four-vectors u_α, p_α and F_α are transformed from one reference frame to another according to transformation equations 5-51.
A four-vector differs from an ordinary vector in one important respect. The norm of a four-vector can be zero and yet the four-vector can have non-zero components. The norm p^2 of a photon's four-momentum is zero due to the fact that $m_0 = 0$ (see Eq.5-65) and yet the photon can have momentum and energy (see Eq.5-66).
In this short introductory course of special theory of relativity we did not study relativistic electrodynamics dealing with the motion of charged particles. The four vectors are extremely useful namely for the study of this topic. As an illustrative example we rewrite the conservation of electric charge law

$$\operatorname{div}\mathbf{j} + \frac{\partial\rho}{\partial t} = 0, \tag{5-68}$$

into four-dimensional form.

We remember that the current density $\mathbf{j} = \rho \mathbf{v}$. Using Eq.5-46 we can write Eq.5-66 as

$$\frac{\partial \rho v_1}{\partial x_1} + \frac{\partial \rho v_2}{\partial x_2} + \frac{\partial \rho v_3}{\partial x_3} + \frac{\partial \rho}{\partial t} \frac{jc}{jc} = 0. \tag{5-69}$$

We denote components of four-dimensional current density as

$$j_1 = \rho v_1 \qquad j_2 = \rho v_2 \qquad j_3 = \rho v_3 \qquad j_4 = jc\rho. \tag{5-70}$$

Do not confuse the components of four-dimensional current density j_1, j_2, j_3 and j_4 with the imaginary unit j. Taking into account Eq.5-46 we can express the conservation of electric charge law as

$$\frac{\partial j_\alpha}{\partial x_\alpha} = 0. \tag{5-71}$$

This is the relativistic statement of the conservation of electric charge law. It can be easily proved that this equation is also Lorentz invariant. It can be also shown that the Maxwell's equations also do obey the principle of relativity.

5.9 Summing Up

The theory of special relativity is confined to the discussion of physical phenomena relative to inertial reference frames. Of course at speed much less than the speed of light the relativistic formulas reduce to the classical ones. Thus we see that the special theory of relativity accomplished a profound unification in physics: it reconciles the physics of low speed with that of high speeds.

Special theory of relativity is not applicable in non-inertial frames of reference. Einstein strove for ten years to develop a general theory of relativity dealing with non-inertial reference frames. He felt that it should be possible to express the laws of physics in a covariant form for all reference frames, that is the mathematical form of the laws of physics should be the same in all frames of reference whether they are accelerating or not. This is the principle of covariance. The general mathematical equations might reduce to simple mathematical forms in inertial reference frames, and they should be consistent with the theory of special relativity in inertial reference frames. In order to develop the full theory Einstein found it necessary to use tensor analysis. Using his full theory of general relativity he was, for example, able to calculate the gravitational shift of spectral lines, the deflection of light near the Sun and the precession of the perihelion of the orbit of the planet Mercury.

It is beyond the scope of the present textbook to discuss here the general relativity. The interested reader can find its description in original books by Einstein, Born, Foch etc (see references).