Chapter 2
Relativistic Dynamics

2.1 Relativistic Energy and Momentum

In the previous chapter we have seen that a proper extension of the principle of relativity to electromagnetism necessarily implies that the correct transformation laws between two inertial frames are the Lorentz transformations. The price we have to pay, however, is that *the laws of classical mechanics are no longer invariant under changes in the inertial reference frame*.[1] We need therefore to re-examine the basic principles of the Newtonian mechanics and to investigate whether they can be made compatible with Einstein's formulation of the principle of relativity which, together with the principle of the constancy of the speed of light, requires invariance under the Lorentz, rather than the Galileo, transformations.

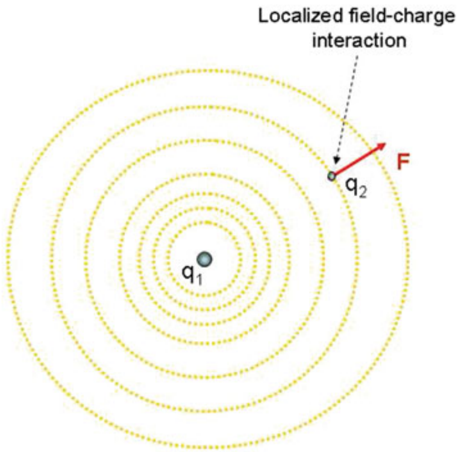
We have indeed learned, form the discussion in the last chapter, that the relativistic kinematics has an important bearing on the very concepts of space and time and, in particular, of *simultaneity*, which are no longer absolute. This fact is incompatible with some of the basic assumptions of classical mechanics. Let us recall that the fundamental force of this theory, the gravitational force, is described as acting at-a-distance. This gives rise to several inconsistencies from the point of view of special relativity:

1. The *instantaneous action* of a body on another, implies the transmission of the interaction at an *infinite velocity*. As we know, no physical signal can propagate with a velocity greater than c. To put it differently, in the action at-a-distance picture, the action of a body A and its effect, consisting in the consequent force applied to B, are simultaneous events localized at different points (corresponding to the positions of A and B respectively). Since simultaneity, in relativistic kinematics, is relative to the reference frame, there will in general exist an observer with respect to which the two events are no longer simultaneous, or in which the force is even seen to act on B before A exerts it, that is before A "knows" about B;

[1] Here by *classical mechanics* we refer to the Newtonian theory.

R. D'Auria and M. Trigiante, *From Special Relativity to Feynman Diagrams*,
UNITEXT for Physics, DOI 10.1007/978-3-319-22014-7_2

Fig. 2.1 Action by contact



2. By the same token also Newton's second law should be revisited. This equation indeed relates the acceleration of a point-mass to the total force exerted on it by all the other bodies, which is given by the sum of the individual forces taken at the same instant, that is *simultaneously*. These forces in turn will depend on the distances between the interacting objects. According to relativistic kinematics both simultaneity and spatial distances are relative to the inertial observer and thus, with respect to a different reference frame, the same forces will appear to be exerted at different times and distances.

The previous considerations imply that a proper formulation of mechanics (and in particular of dynamics) has to be given in terms of *localized interactions*, that is in terms of an interaction which takes place only when the two interacting parts are *in contact* and which is then localized in a certain point. This is in fact what happens when two point-charges interact through the electromagnetic field. The interaction is no longer represented as an action at-a-distance between the two charges but as mediated by the electromagnetic field, and can be divided into two moments (see the Fig. 2.1):

(a) a charge q_1 generates an electromagnetic field;
(b) The field, which is a physical quantity defined everywhere in space, propagates until it reaches the charge q_2 located at some point and acts on it by means of a force (the Lorentz force).

This mechanism is apparent when one of the two charges (say q_1) is moving at a very high speed. One then observes that the information about the position of the moving charge is transmitted to q_2 at the speed of light through the electromagnetic field, causing the force acting on it to be adjusted accordingly with a characteristic delay which depends on the distance between the two charges. In this *action-by-contact* picture the interacting parts are three instead of just two: the two charges and the field. The force acting on q_2 is the effect of the action of the field generated by q_1

on q_2. This implies that the action and the resulting force occur at the same time and place (the position of q_2) and this property is now Lorentz-invariant. Indeed if

$$\Delta t = |\Delta \mathbf{x}| = 0, \tag{2.1}$$

in a given frame, using the Lorentz transformations (1.58)–(1.61), we also have $\Delta t' = |\Delta \mathbf{x}'| = 0$ in any other frame. Thus *the action-by-contact representation is consistent with the principles of relativity and causality.*

As for the electromagnetic interaction, we would also expect the gravitational one to be mediated by a gravitational field. However, as we have mentioned earlier, a correct treatment of the gravitational interaction requires considering non-inertial frames of reference which goes beyond the framework of special relativity. In order to discuss how classical mechanics should be generalized in order to be compatible with Lorentz transformations (relativistic mechanics), we shall therefore refrain from considering gravitational interactions.

Even in classical mechanics we can consider processes in which the interaction is localized in space and time, so that the locality condition (2.1) is satisfied and we can avoid the inconsistencies discussed above, related to Newton's second law. These are typically *collisions* in which two or more particles interact for a very short time and in a very small region of space. Since the strength of the interaction is much higher than that of any other external force acting on the particles, the system can be regarded as isolated, so that the total linear momentum is conserved, and its initial and final states are described by free particles. Let us focus on this kind of processes in order to illustrate how one of the fundamental laws of classical mechanics, the conservation of linear momentum, can be made consistent with the principle of relativity, as implemented by the Lorentz transformations.

We shall first show that, *if we insist in defining the mass as independent of the velocity, then the conservation of momentum cannot hold in any reference frame*, thus violating the principle of relativity.[2]

Let us consider a simple process in which a mass m explodes into two fragments of masses $m_1 = m_2 = m/2$ (or equivalently a particle of mass m decays into two particles of equal masses, see Fig. 2.2). We shall assume the conservation of linear momentum to hold in the frame S in which the exploding mass is at rest:

$$\mathbf{v}_m = \mathbf{0} = \frac{m}{2}\,\mathbf{v}_1 + \frac{m}{2}\,\mathbf{v}_2 \quad \Rightarrow \quad \mathbf{v}_1 = -\mathbf{v}_2.$$

For the sake of simplicity we take the x-axis along the common direction of motion of the particles after the collision, so that $v_{1(y,z)} = v_{2(y,z)} = 0$. Let us now check whether the conservation of linear momentum also holds in a different frame S'. We choose S' to be the rest frame of fragment 1, which moves along the positive x-direction at a constant speed $V = v_{1(x)} \equiv v_1$ relative to S, and let the explosion

[2]Here and in the rest of this chapter, when referring to the conservation of the total linear momentum of an isolated system of particles, we shall often omit to specify that we consider the *total* momentum and that the system is *isolated*, regarding this as understood.

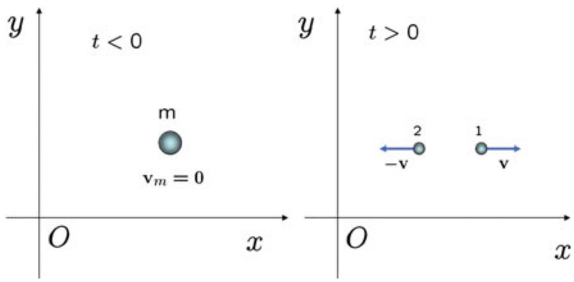


Fig. 2.2 Decay of a particle into two particles of equal masses

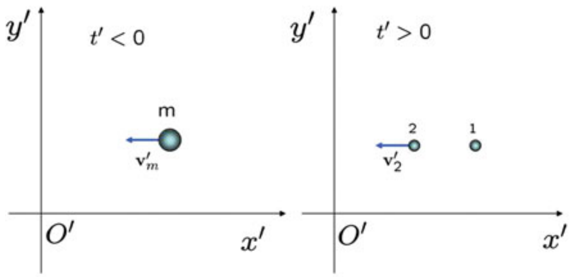


Fig. 2.3 Same decay in the rest frame of particle 1

occur at the instant $t = t' = 0$, see Fig. 2.3. In the frame S', the velocity of the mass m before the explosion is obtained by applying the relativistic composition law for velocities (1.76):

$$v'_{(m)x} = \frac{0 - v_1}{1 - \frac{0 \cdot v_1}{c^2}} = -v_1 = -V, \quad v'_{(m)y,z} = 0.$$

Analogously, after the explosion, the velocities of the fragments in S' are given by:

$$v'_1 \equiv v'_{1x} = 0, \quad v'_2 \equiv v'_{2x} = \frac{-v_1 - v_1}{1 + \frac{v_1^2}{c^2}}, \quad v'_{2y} = v'_{2z} = 0.$$

Having computed the velocities in S' we may readily check whether the conservation of linear momentum holds in this frame. It is sufficient to consider the components of the linear momenta along the common axis $x = x'$; before and after the explosion the total momenta in S' are given respectively by:

$$P'_{in} = m\, v'_{(m)} = -m\, v_1, \tag{2.2}$$

$$P'_{fin} = \frac{m}{2}\, v'_1 + \frac{m}{2}\, v'_{2x} = \frac{m}{2}\, v'_2 = -\frac{m\, v_1}{1 + \frac{v_1^2}{c^2}}. \tag{2.3}$$

Since

$$m\,v_1 \neq \frac{m\,v_1}{1+\frac{v_1^2}{c^2}}, \tag{2.4}$$

we conclude that, *in S' the total momentum is not conserved*. Or better, *the principle of conservation of linear momentum (as defined in classical mechanics) is not covariant under Lorentz transformations* thereby violating the principle of relativity. As such it cannot be taken as a founding principle of the new mechanics. It is clear that, just as the principle of relativity cannot be avoided in any physical theory, it would also be extremely unsatisfactory to give up the conservation of linear momentum; in the absence of it we would indeed be deprived of an important guiding principle for building up a theory of mechanics. To remediate this apparent shortcoming, it is important to trace back, in the above example, the origin of the non-conservation of the total momentum.

For this purpose we note the presence of the irksome factor $1+\frac{v^2}{c^2}$ on the right hand side of the inequality (2.4), which reduces to 1 in the non relativistic limit. This factor derives from the peculiar form of the composition law of velocities, which, in turn, originates from the non-invariance of time intervals under Lorentz transformations, namely:

$$dt' = \gamma(V)\left(dt - \frac{V}{c^2}\,dx\right) = \gamma(V)\,dt\left(1 - \frac{V\,v_x}{c^2}\right).$$

Thus we see that the non-trivial transformation property of dt is at the origin of the apparent failure of the conservation of momentum.

The same fact, however, gives us the clue to the solution of our problem: If we indeed replace, in the definition of the linear momentum $\mathbf{p}$ of a particle, the non-invariant time interval dt with the *proper time* $d\tau$, which is *invariant* under a change in the inertial frame, we may hope to have a conservation law of momentum compatible with the Lorentz transformations.

Let us then try to define the *relativistic linear momentum* of a particle as follows:

$$\mathbf{p} = m\frac{d\mathbf{x}}{d\tau}. \tag{2.5}$$

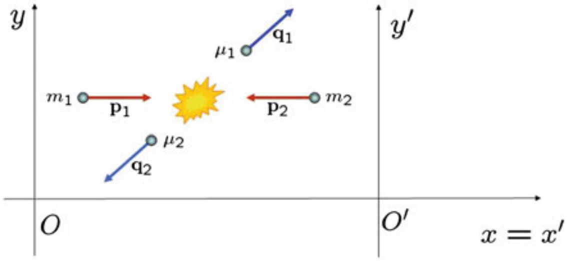
Recalling the relation between dt and $d\tau$, given by Eq. (1.76), and equations below, of the previous chapter,

$$d\tau = \frac{1}{\gamma(v)}dt = \sqrt{1-\frac{v^2}{c^2}}\,dt, \tag{2.6}$$

we may write:

$$\mathbf{p} = m\,\frac{d\mathbf{x}}{d\tau} = m\,\gamma(v)\,\frac{d\mathbf{x}}{dt} = m(v)\,\mathbf{v}, \tag{2.7}$$

Fig. 2.4 Collision



where

$$m(v) \equiv m\,\gamma(v) = \frac{m}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{2.8}$$

Note that the new definition of the relativistic momentum, (2.7), can be obtained from the classical one by *replacing the constant (classical) mass* m, with the velocity-dependent quantity $m(v)$, called the *relativistic mass*, so that the classical mass m coincides with the relativistic one only when the body is at rest: $m = m(v = 0)$. The mass m is then called *the rest mass* of the particle.

Let us now show that the conservation law of linear momentum is relativistic, provided we use (2.7) as the definition of the linear momentum of a particle. To prove the validity of this principle we would need to consider the most general process of interaction within an isolated system. For the sake of simplicity, we shall still restrict ourselves to collision processes, in order to deal with localized interactions, between two particles only. Consider then a process in which two particles of rest masses $m_1,\ m_2$ and linear momenta $\mathbf{p}_1, \mathbf{p}_2$ collide and two new particles are produced with rest masses and momenta $\mu_1,\ \mu_2$ and $\mathbf{q}_1,\ \mathbf{q}_2$, respectively, see Fig. 2.4.

We assume that, in a given frame S, the conservation of total linear momentum holds:

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{q}_1 + \mathbf{q}_2. \tag{2.9}$$

The above equation, using the definitions (2.7), can be rewritten in the following equivalent forms

$$\begin{aligned} m_1(v_1)\mathbf{v}_1 + m_2(v_2)\mathbf{v}_2 &= \mu_1(u_1)\mathbf{u}_1 + \mu_2(u_2)\mathbf{u}_1, \\ m_1\frac{d\mathbf{x}_1}{d\tau_1} + m_1\frac{d\mathbf{x}_1}{d\tau_2} &= \mu_1\frac{d\tilde{\mathbf{x}}_1}{d\tilde{\tau}_1} + \mu_2\frac{d\tilde{\mathbf{x}}_2}{d\tilde{\tau}_2}, \end{aligned} \tag{2.10}$$

where we have marked with a tilde the quantities referring to the final state.

For the purpose of writing the conservation law in a new reference frame, we shall find it more useful to work with the second of Eqs. (2.10).

Let us consider now the process from a new frame S' moving with respect to S at constant speed, in the standard configuration. The two descriptions are related by a Lorentz transformation. In particular, if we apply the Lorentz transformation to

Eq. (2.10), we note first of all that the components of the same equation along the y- and z-axes do not change their form since the lengths along these directions are Lorentz-invariant ($dy' = dy, dz' = dz$), as well as the rest masses m_i, μ_i and the proper time intervals $d\tau_i$, $d\tilde{\tau}_i$. We can therefore restrict to the only component of Eq. (2.10) along the x-axis and prove that

$$m_1 \frac{dx_1}{d\tau_1} + m_1 \frac{dx_2}{d\tau_2} = \mu_1 \frac{d\tilde{x}_1}{d\tilde{\tau}_1} + \mu_2 \frac{d\tilde{x}_2}{d\tilde{\tau}_2}, \tag{2.11}$$

has the same form in the frame S', namely that it is *covariant* under a standard Lorentz transformation. This is readily done by transforming the differentials $dx_i, d\tilde{x}_i$ in Eq. (2.10), according to the inverse of transformation (1.58):

$$\sum_{i=1}^{2} m_i \left(\frac{dx_i'}{d\tau_i} + V \frac{dt'}{d\tau_i} \right) \gamma(V) = \sum_{i=1}^{2} \mu_i \left(\frac{d\tilde{x}_i'}{d\tilde{\tau}_i} + V \frac{dt'}{d\tilde{\tau}_i} \right) \gamma(V).$$

Let us now perform, using Eq. (2.7), the following replacement

$$\frac{d}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt} = \gamma(v)\frac{d}{dt},$$

v being the velocity of the particle, so that, recalling the definition of the relativistic momentum, (2.11) takes the following form[3]:

$$\left(\sum_i p_{xi}' - \sum_i q_{xi}' \right) = \gamma(V)\, V \sum_i \left(m_i \gamma(v_i') - \mu_i \gamma(u_i') \right),$$

where v_i' and u_i', as usual, denote the velocities of the particles before and after the collision in the frame S'. The above relation can also be written in vector form as follows:

$$\left(\sum_i \mathbf{p}_i' - \sum_i \mathbf{q}_i' \right) \propto \sum_i \left(m_i(v_i') - \mu_i(u_i'),\ 0,\ 0 \right). \tag{2.12}$$

[3]Note that $\gamma(V) = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}$ is the relativistic factor associated with the motion of S' relative to S, while $\gamma(v_i') = \frac{1}{\sqrt{1-\frac{v_i'^2}{c^2}}}$ and $\gamma(u_i') = \frac{1}{\sqrt{1-\frac{u_i'^2}{c^2}}}$ are the relativistic factors depending on the velocities of each particle and relate the time dt' in S' to the proper times $d\tau_i, d\tilde{\tau}_i$ referred to the rest-frames of the various particles, according to

$$\begin{cases} d\tau_i = \sqrt{1 - \frac{v_i'^2}{c^2}} dt' \\ d\tilde{\tau}_i = \sqrt{1 - \frac{u_i'^2}{c^2}}\, dt'. \end{cases}$$

Since the right hand side contains only the difference between the sum of the *relativistic masses* before and after the collision, it follows that, in order for the conservation of linear momentum to hold in the new reference frame

$$\sum_i \mathbf{p}'_i = \sum_i \mathbf{q}'_i, \tag{2.13}$$

we must have:

$$\sum_i m_i(v'_i) = \sum_i \mu_i(u'_i), \tag{2.14}$$

that is the total relativistic mass must be conserved. From this analysis we can conclude that:

Given the new definition of linear momentum, (2.7)*, the conservation of momentum is consistent with the principle of relativity, i.e. covariant under Lorentz transformations, if and only if the total relativistic mass is also conserved.*

Let us emphasize the deep analogy between our present conclusion and the analogous result obtained when studying the covariance of the conservation law of momentum under Galilean transformations in Newtonian mechanics (see Sect. 1.1.2).

2.1.1 Energy and Mass

We have seen that the concept of force as an action at a distance on a given particle looses its meaning in a relativistic theory. However nothing prevents us from *defining* the force acting on a particle as the time derivative of its relativistic momentum:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}. \tag{2.15}$$

Recalling the definition of $\mathbf{p}$, namely, $\mathbf{p} = m(v)\,\mathbf{v}$, we find:

$$\mathbf{F} = \frac{d}{dt}\,(m(v)\mathbf{v}) = \frac{dm(v)}{dt}\,\mathbf{v} + m(v)\frac{d\mathbf{v}}{dt},$$

Note that $\mathbf{F}$ is in general no longer proportional to the acceleration $\mathbf{a} = \frac{d\mathbf{v}}{dt}$. Writing $\mathbf{v} = v\,\mathbf{u}$, where $\mathbf{u}$ is the unit vector in the direction of motion, we obtain

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \left(\frac{dv}{dt}\right)\mathbf{u} + \frac{v^2}{\rho}\,\mathbf{n},$$

where, as is well known, the unit vector $\mathbf{n}$ is normal to $\mathbf{u}$ and oriented towards the concavity of the trajectory, ρ being the radius of curvature.

Computing the time derivative of the relativistic mass we find

$$\frac{dm(v)}{dt} = m\frac{d}{dt}\left(\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\right) = \frac{m}{\left(1-\frac{v^2}{c^2}\right)^{3/2}}\frac{v}{c^2}\frac{dv}{dt} = \frac{m(v)}{c^2}\frac{v}{1-\frac{v^2}{c^2}}\frac{dv}{dt},$$

so that:

$$\mathbf{F} = m(v)\left(\frac{1}{1-\frac{v^2}{c^2}}\frac{v^2}{c^2}\frac{dv}{dt}+\frac{dv}{dt}\right)\mathbf{u} + m(v)\frac{v^2}{\rho}\mathbf{n} = \frac{m(v)}{1-\frac{v^2}{c^2}}\frac{dv}{dt}\mathbf{u} + m(v)\frac{v^2}{\rho}\mathbf{n}.$$

We are now ready to determine the relativistic expression for the kinetic energy of a particle by computing the work done by the total force $\mathbf{F}$ acting on it. For an infinitesimal displacement $d\mathbf{x} = \mathbf{v}\,dt$ along the trajectory, the work reads:

$$dW = \mathbf{F}\cdot d\mathbf{x} = \mathbf{F}\cdot\mathbf{v}\,dt = \frac{m(v)}{1-\frac{v^2}{c^2}}\,v\,\frac{dv}{dt}\,dt = \frac{1}{2}\frac{m(v)}{1-\frac{v^2}{c^2}}\,d(v^2).$$

Integrating along the trajectory Γ (and changing the integration variable into $x = 1-\frac{v^2}{c^2}$), we easily find:

$$W = \int_\Gamma \mathbf{F}\cdot d\mathbf{x} = \int_\Gamma \frac{1}{2}\frac{m}{\left(1-\frac{v^2}{c^2}\right)^{3/2}}dv^2 = -\frac{c^2}{2}m\int\frac{dx}{x^{3/2}} = mc^2\left(\frac{1}{x^{1/2}}\right)$$
$$= \frac{m}{\sqrt{1-\frac{v^2}{c^2}}}c^2 + \text{const.} = m(v)c^2 + \text{const.}$$

If we define the kinetic energy, as in the classical case, to be zero when the particle is at rest, then the constant is determined to be $-m(0)c^2$, so that, the kinetic energy E_k acquired by the particle will be given by:

$$E_k(v) = m(v)\,c^2 - m\,c^2, \tag{2.16}$$

where, from now on, $m = m(0)$ will always denote the rest mass. Note that in the non-relativistic limit $v^2/c^2 \ll 1$, we retrieve the Newtonian result:

$$E_k(v) = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} - mc^2 \simeq mc^2\left(1+\frac{1}{2}\frac{v^2}{c^2}\right) - mc^2 = \frac{1}{2}mv^2. \tag{2.17}$$

where we have neglected terms of order $O(v^4/c^4)$.

Let us define the *total energy* of a body as:

$$E = m(v)\,c^2, \tag{2.18}$$

The kinetic energy is then expressed, in (2.16), as the difference between the total energy and the *rest energy*, which is the amount of energy a mass possesses when it is at rest:

$$E_{rest} = E(v = 0) = m\,c^2. \tag{2.19}$$

To motivate the definition of the total energy of a particle given in (2.18), we prove that the total energy in a collision process, defined as the sum of the total energies of each colliding particle, is always conserved. This immediately follows from the conservation of the total (relativistic) mass, which we have shown to be a necessary requirement for the conservation law of momentum to be covariant. Indeed by multiplying both sides of

$$m_1(v_1) + m_2(v_2) + \cdots + m_k(v_k) = \text{const.}$$

by c^2 and using the definition (2.18), we find

$$E_1(v_1) + E_2(v_2) + \cdots + E_k(v_k) = \text{const.}$$

This fact has no correspondence in classical mechanics where we know that, as opposed to the total linear momentum, which is always conserved in collision processes, the conservation of mechanical energy only holds in elastic collisions. This apparent clash between the classical and the relativistic laws of energy conservation is obviously a consequence of the fact that the rest energy can be transformed into other forms of energy, like kinetic energy, etc. We can give a clear illustration of this by considering again the collision of two particles with rest masses m_1, m_2 and velocities $\mathbf{v}_1$ e $\mathbf{v}_2$. Suppose that the collision is perfectly inelastic, so that the two particles stick together into a single one of rest mass M. It is convenient to describe the process in the center of mass frame, in which the final particle is at rest. Let us first describe the collision in the context of Newtonian mechanics. The conservation of momentum reads:

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{P} = 0,$$

or, equivalently

$$m_1\mathbf{v}_1 + m_2\mathbf{v}_2 = 0.$$

Moreover conservation of the *classical mass* is also assumed.

$$M = m_1 + m_2. \tag{2.20}$$

The initial and final mechanical (kinetic) energies are however different since

$$E_k^i = m_1 \frac{v_1^2}{2} + m_2 \frac{v_2^2}{2}, \quad E_k^f = 0.$$

and thus

$$\Delta E_k = E_k^f - E_k^i = -\left(m_1 \frac{v_1^2}{2} + m_2 \frac{v_2^2}{2}\right) \neq 0.$$

In Newtonian physics, the interpretation of this result is that the kinetic energy of the particles in the initial state is not conserved, while, from the thermodynamical point of view it has been converted into heat, increasing the thermal energy of the final body, that is the disordered kinetic energy of the constituent molecules.

Let us now describe the same process from the relativistic point of view.

Conservation of momentum and energy give the following two equations:

$$\begin{aligned} m_1 \gamma(v_1)\mathbf{v}_1 + m_2 \gamma(v_2)\mathbf{v}_2 &= 0 \\ m_1(v_1)c^2 + m_2(v_2)c^2 &= M(0)c^2 \end{aligned}$$

where we have set $M(v = 0) = M(0)$. Using Eq. (2.16) to separate the rest masses from the (relativistic) kinetic energies, we obtain

$$E_k(v_1) + m_1 c^2 + E_k(v_2) + m_2 c^2 = E_k^f(0) + M(0)c^2. \tag{2.21}$$

where on the right hand side $E_k^f(0) \equiv E_k^f(v = 0) = 0$. It follows[4]

$$c^2 \Delta M(0) \equiv c^2(M(0) - m_1 - m_2) = -(0 - E_k(v_1) - E_k(v_2)) = -\Delta E_k. \tag{2.22}$$

From the above relation we recognize that the loss of kinetic energy has been transformed in an increase of the final rest mass $M = M(0)$; thus M is not the sum of the rest masses of the initial particles (as it was instead assumed in the classical case, see Eq. (2.20)).

If we consider the inverse process in which a particle of rest mass M decays, in its rest frame, into two particles of rest masses m_1 and m_2, we see that part of the initial rest mass is now converted into the kinetic energy of the decay products. The importance of this effect obviously depends on the size of the ratio (v^2/c^2).

These examples illustrate an important implication of relativistic dynamics: The rest mass m of an object can be regarded as a form of energy, the rest energy $m\,c^2$, which can be converted into other forms of energy (kinetic, potential, thermal etc.). Let us illustrate this property in an other example. Consider a body of mass M at some given temperature: M will be given by the sum of the *relativistic masses* of its constituent molecules, and its temperature is related to their thermal motion. If we

[4]Note that at order $O(v^2/c^2)$ Eq. (2.21) can be written $c^2 \Delta M(0) = (\frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2)$.

now transfer an amount of energy E, in the form of heat, to the body, the total kinetic energy of its molecules will increase by E, thereby implying an increase in the mass by E/c^2, $M \to M + E/c^2$. Since all forms of energy can be transformed into one another and in particular into heat, we see that we can associate an equivalent amount of energy E with any mass, and, in particular, with the rest mass $m = m(0)$. Vice versa to each form of energy there corresponds an equivalent amount of mass given by $m(v) = E/c^2$.

The *equivalence between mass ed energy*, which is expressed by Eq. (2.18) or, equivalently, by

$$\Delta E = \Delta m(v)\, c^2, \tag{2.23}$$

is one of the major results of Einstein's theory of relativity. As a consequence, for a system made of interacting parts we define the total energy as:

$$E_{tot} = E_0 + E_k + U + \ldots, \tag{2.24}$$

where the sum is made over all the forms of energy which are present in the system: total rest energy, kinetic energy, potential energy and so on.

As a further example, let us consider a *bound system*. By definition a bound system is a system of interacting bodies such that the sum of the kinetic and potential energies is negative:

$$E_k + U < 0, \tag{2.25}$$

provided we fix the potential energy U to be zero when all the components are at infinite distance from each other and thus non-interacting: $U_\infty = 0$. If we think of the bound system as a single particle of rest mass M, in its rest frame we can write its energy $E_0 = M\, c^2$ as the total energy of the system,namely as the sum of the rest energies of its constituents and their total kinetic and potential energies, according to Eq. (2.24):

$$E_0 = M\, c^2 = \sum_i E_{0\,i} + E_k + U = \sum_i m_i\, c^2 + E_k + U, \tag{2.26}$$

m_i being the rest masses of the constituent particles. Equations (2.25) and (2.26) imply that, in order to disassemble the system bringing its elementary parts to infinite distances from one another (non-interacting configuration), we should supply it with an amount of energy (called the *binding energy* of the system) given by

$$\Delta E = -(E_k + U) > 0.$$

Note that, being $E_k + U$ a negative quantity, from Eq. (2.26) it immediately follows that the rest mass of the bound state is smaller than the sum of the rest masses of its constituents

$$M = \sum_i m_i - \frac{\Delta E}{c^2} < \sum_i m_i, \tag{2.27}$$

the "missing" rest mass being the equivalent in mass of the binding energy, as it follows from Eq. (2.26)

$$\Delta M \equiv \sum_i m_i - M = \frac{\Delta E}{c^2}. \tag{2.28}$$

Therefore when a bound state of two or more particles is formed starting from a non-interacting configuration, the system looses part of its total rest mass which, being the total energy conserved, is converted into an equivalent amount of energy $\Delta E = \Delta M\, c^2$ and released as, for instance, radiation.

An example of bound state is the hydrogen atom. It consists of a positively charged proton and a negatively charged electron, the two being bound together by the electric force. The rest masses of the two particles are respectively:

$$m_p \cong 938.3 \text{ MeV}/c^2; \qquad m_e \cong 0.5 \text{ MeV}/c^2,$$

where, taking into account the equivalence between mass and energy, we have used for the masses the unit MeV$/c^2$.[5] The corresponding binding energy

$$\Delta E = 1\mathbb{R}_y \cong 13.5\,\text{eV},$$

is called a *Rydberg*. Since the rest energy of the hydrogen atom is

$$M\,c^2 = m_e\,c^2 + m_p\,c^2 - \Delta E = \left(938.3 \times 10^6 + 0.5 \times 10^6 - 13.5\right) \text{ eV},$$

it follows that

$$\frac{\Sigma_i m^i\,c^2 - M\,c^2}{\Sigma_i m^i\,c^2} = \frac{\Delta M}{m_e + m_p} = \frac{13.5}{938.8 \times 10^6} \cong 10^{-8}. \tag{2.29}$$

Thus we see that, in this case, where the force in play is the electric one, the rate of change in rest mass, $\Delta M/M$, is quite negligible.

---

[5] We recall that 1 MeV $= 10^6$ eV, where 1 eV is the energy acquired by an electron (whose charge is e $\cong 1.6 \times 10^{-19}$ C) crossing an electric potential difference of 1 V:

$$1\,\text{eV} = 1.6 \times 10^{-19}\ \text{J}.$$

Another commonly used unit, when considering energy exchanges in atomic processes, is the atomic mass unit u, that is defined as $1/12$ the rest mass M_C of the isotope ${}^{12}C$ of the carbon atom at rest; this unit is more or less the proton mass. Precisely we have: $1\,u = 1.660\,538\,782(83) \times 10^{-24}$ g $= \frac{1}{N_A}$ g, where N_A is the Avogadro number. Taking into account the equivalence mass-energy we also have

$$1\,u = \frac{1}{12}\,M_C \simeq 931.494\,\text{MeV}/c^2.$$

Let us now consider a two-body system bound by the *nuclear force*, The ratio of the strength of the nuclear force to that of the electric one is of order 10^5. An example is the deuteron system which is a bound state of a proton and a neutron. In this case we may expect a much larger binding energy and, consequently, a greater rest mass variation. Using the values of the proton and neutron masses,

$$\begin{aligned} c^2\, m_p &\cong 938.272\,\mathrm{MeV} \simeq 1.00728\, u, \\ c^2\, m_n &\cong 939.566\,\mathrm{MeV} \simeq 1.00867\, u, \\ \Delta E &\simeq 2.225\,\mathrm{MeV}, \end{aligned}$$

it turns out that the corresponding loss of rest mass is:

$$\frac{-\Delta M}{m_p + m_n} = \frac{2.225}{1877.838} = 1.18 \times 10^{-3},$$

that is, five orders of magnitudes greater than in the case of the hydrogen atom. This missing rest mass (times the square of the speed of light) results in an amount of energy which is released when the bound state is created. A similar mass defect is present in all atomic nuclei. In fact, as the reader can easily verify, the atomic mass of an atom, which can be read off the Mendeleev table, is always smaller than the sum of the masses of the protons and neutrons entering the corresponding nuclei, since they form a bound state.

2.1.2 Nuclear Fusion and the Energy of a Star

Taking into account that life on earth depends almost exclusively on the energy released by the sun, it is of outmost importance to realize that the source of such energy is the continuous conversion of the solar rest mass into radiation energy and heat that we receive on earth, through the so-called *nuclear fusion*; just as for the reaction discussed above, leading to the creation of the deuteron, nuclear fusion essentially amounts to the formation of a bound state of nucleons (protons and neutrons) with a consequent reduction of rest mass which is released in the form of energy (radiation). The fact that the solar energy, or more generally, the energy of a star, could not originate from chemical reactions, can be inferred from the astronomical observation that the mean life of a typical star, like the sun, is of the order of 10^9–10^{10} years. If the energy released by the sun were of chemical origin, one can calculate that the mean life of the sun would not exceed 10^5–10^6 years. It is only through the conversion of mass into energy, explained by the theory of special relativity, that the lifetime of stars can be fully explained in relation to their energy emission.

Without entering into a detailed description of the sequences of nuclear processes taking place in the core of a burning star (which also depend on the mass of the star), we limit ourselves to give a qualitative description of the essential phenomenon.

We recall that after the formation of a star, an enormous gravitational pressure is generated in its interior, so that the internal temperature increases to typical values of 10^6–10^7 K. At such temperatures nuclear fusion reactions begin to take place, since the average kinetic energy of nucleons is large enough to overcome the repulsive (electrostatic) potential barrier separating them. At sufficiently short distances, the interaction between nucleons is dominated by the attractive nuclear force and nucleon bound states can form. The fundamental reaction essentially involves four protons which give rise, after intermediate processes, to a nucleus of Helium, ${}^4_2He$, together with two *positrons and neutrinos*:

$$4 \times {}^1_1H \rightarrow {}^4_2He + 2e^+ + 2\nu_e. \tag{2.30}$$

where e^+ denotes the positron (the anti-particle of an electron) and ν_e the (electronic) neutrino, their masses being respectively: $m_{e^+} = m_{e^-} \simeq 0.5\,\text{MeV}$, $m_{\nu_e} \simeq 0$.
(Note that ionized hydrogen, that is protons, comprise most of the actual content of a star.)

The reaction (2.30) is the aforementioned nuclear fusion taking place in the interior of a typical star. To evaluate the mass reduction involved in this reaction we use the value of the mass of a ${}^4He$ nucleus, and obtain:

$$\Delta M \cong 0.0283\,u = 0.0283 \times 931.494\,\text{MeV}/c^2 \simeq 26.36\,\text{MeV}/c^2.$$

This implies that every time a nucleus of ${}^4He$ is formed out four protons, an amount of energy of about 26.36 MeV is released.

Consider now the fusion of 1 kg of ionized hydrogen. Since 1 mole of ${}^1_1H$, weighting about 1 g, contains $N_A \simeq 6.023 \times 10^{23}$ (Avogadro's number) particles, there will be a total of $\sim 1.5 \times 10^{26}$ reactions described by (2.30), resulting in an energy release of[6]:

$$\Delta E(1\,\text{kg}) = 26.36 \times 1.5 \times 10^{26}\,\text{MeV} \simeq 3.97 \times 10^{27}\,\text{MeV} \approx 6.35 \times 10^{14}\,\text{J}.$$

On the other hand, we know that a star like the sun fuses H_1^1 at a rate of about $5.64 \times 10^{11}\,\text{kg}\,\text{s}^{-1}$, the total energy released every second by our star amounts approximately to:

$$\frac{\Delta E}{\Delta t} = 6.35 \times 10^{14} \times 5.64 \times 10^{11} \approx 3.58 \times 10^{26}\,\text{J}\,\text{s}^{-1},$$

This implies a reduction of the solar mass at a rate of:

[6]Note that if we had a chemical reaction instead of a nuclear one, involving just the electrons of two hydrogen atoms ($H + H \rightarrow H_2$) we would obtain an energy release of $E \simeq 2 \times 10^6$ J, which is eight orders of magnitude smaller.

$$\frac{\Delta m}{\Delta t} = \frac{1}{c^2}\frac{\Delta E}{\Delta t} \cong 3.98 \times 10^9\,\mathrm{kg\,s^{-1}}.$$

Since the energy emitted over one year ($=3.2 \times 10^7$ s) is $\Delta E_{year} \simeq 1.1 \times 10^{34}$ J, the corresponding mass lost each year by our sun is $\Delta M_{year} \simeq 1.3 \times 10^{17}$ kg. If this loss of mass would continue indefinitely,[7] using the present value of the solar mass, $M_{\odot} \simeq 1.9 \times 10^{30}$ kg, its mean life can be roughly estimated to be of the order of

$$T = \frac{M_{\odot}}{\Delta M}(\text{years}) \simeq 1.5 \times 10^{13}\ \text{years}.$$

2.2 Space-Time and Four-Vectors

It is useful at this point to introduce a mathematical set up where all the kinematic quantities introduced until now and their transformation properties have a natural and transparent interpretation.

To summarize our results so far, the energy and momentum of a particle of rest mass m moving at velocity $\mathbf{v}$ in a given frame S, are defined as:

- *energy* : $E = m(v)c^2 = m\gamma(v)c^2 = m\frac{dt}{d\tau}c^2$,
- *momentum*: $\mathbf{p} = m\frac{d\mathbf{x}}{d\tau} = m(v)\mathbf{v}$,

where $\mathbf{v} = \frac{d\mathbf{x}}{dt}$, and $m(v) = \frac{m}{\sqrt{1-\frac{v^2}{c^2}}} = m\,\frac{dt}{d\tau}$.

From the above definitions we immediately realize that the four quantities:

$$\left(\frac{E}{c}, \mathbf{p}\right) \equiv \left(m\frac{dt}{d\tau}, m\frac{d\mathbf{x}}{d\tau}\right), \tag{2.31}$$

transform exactly as $(c\,dt,\ d\mathbf{x})$ under a Lorentz transformation, since both m and $d\tau$ are invariant. Thus, using a standard configuration for the two frames in relative motion with velocity V, we may readily compute the transformation law of $E,\ \mathbf{p}$:

$$\begin{aligned} p'_x &= m\frac{dx'}{d\tau} = m\gamma(V)\frac{dx - V\,dt}{d\tau} = \gamma(V)\left(m\frac{dx}{d\tau} - V\,m\frac{dt}{d\tau}\right), \\ &= \gamma(V)\left(p_x - V\,\frac{E}{c^2}\right). \end{aligned} \tag{2.32}$$

[7] This does not happen however, because the nuclear fusion of hydrogen ceases when there is no more hydrogen, and after that new reactions and astrophysical phenomena begin to take place.

where we have used that $m\, dt/d\tau = m\, \gamma(v) = E/c^2$. Furthermore we also have

$$p'_y = p_y, \tag{2.33}$$

$$p'_z = p_z, \tag{2.34}$$

$$\frac{E'}{c} = mc\,\frac{dt'}{d\tau} = m\,\gamma(V)\frac{c\,dt - \frac{V}{c}\,dx}{d\tau} = \gamma(V)\left(\frac{E}{c} - \frac{V}{c}\,p_x\right), \tag{2.35}$$

where we have used the property that the proper time interval, as defined in Eq. (1.76), is Lorentz-invariant: $d\tau' = d\tau$. Comparing the transformation laws for the time and spatial coordinates with those for energy and momentum, given by the Eqs. (2.32)–(2.35), we realize that, given the correspondences $(p_x, p_y, p_z) \rightarrow (dx, dy, dz)$ and $E/c \rightarrow c\,dt$, they are identical:

$$\begin{cases} p'_x = \gamma(V)\left(p_x - V\frac{E}{c^2}\right) \\ p'_y = p_y \\ p'_z = p_z \\ \frac{E'}{c} = \gamma(V)\left(\frac{E}{c} - \frac{V}{c}\,p_x\right) \end{cases} \leftrightarrow \begin{cases} dx' &= \gamma(V)\,(dx - V dt) \\ dy' &= dy \\ dz' &= dz \\ c\,dt' &= \gamma(V)\left(c\,dt - \frac{V}{c}\,dx\right) \end{cases} \tag{2.36}$$

We now recall the expression of the Lorentz-invariant proper time interval, as defined in Eq. (1.76):

$$d\tau^2 = dt^2 - \frac{1}{c^2}|d\mathbf{x}|^{\mathbf{2}}. \tag{2.37}$$

From the above correspondence it follows that the analogous quantity

$$\frac{E^2}{c^2} - |\mathbf{p}|^2 = m^2\gamma(v)^2\,c^2 - m^2\gamma(v)^2\,v^2 = \frac{m^2c^2}{1 - \frac{v^2}{c^2}}\left(1 - \frac{v^2}{c^2}\right) = m^2c^2,$$

is Lorentz-invariant as well, being simply proportional to the rest mass of the particle.

Note that the relativistic relation between energy and momentum given by

$$\frac{E^2}{c^2} - |\mathbf{p}|^2 = m^2c^2 \quad \Rightarrow \quad E = \sqrt{\mathbf{p}^2\,c^2 + m^2c^4}, \tag{2.38}$$

separating the relativistic kinetic energy from the rest mass, can be rewritten as follows:

$$mc^2 + E_k^{rel.} = mc^2\left(\sqrt{1 + \frac{\mathbf{p}_{rel}^2}{m^2c^2}}\right). \tag{2.39}$$

In the non-relativistic limit, neglecting higher order terms in v^2/c^2, Eq. (2.39) becomes:

$$m\,c^2 + \frac{1}{2}m\,v^2 \simeq m\,c^2\left(1 + \frac{\mathbf{p}_{rel}^2}{2m^2c^2}\right) \;\Rightarrow\; \frac{1}{2}m\,v^2 = \frac{\mathbf{p}_{class.}^2}{2m}, \tag{2.40}$$

in agreement with the standard relation between kinetic energy and momentum in classical mechanics.

2.2.1 Four-Vectors

In the previous chapter we have seen that the time and space coordinates of an *event* may be regarded as coordinates (ct, x, y, z) of a four-dimensional space-time called *Minkowski space*, for which we shall use the following short-hand notation

$$(x^\mu) = (x^0, x^1, x^2, x^3) = (ct, x, y, z); \quad (\mu = 0, 1, 2, 3),$$

The time coordinate $x^0 = ct$ has been defined in such a way that all the four coordinates x^μ share the same dimension. These coordinates can be viewed as the orthogonal components of the position vector of an event relative to the origin-event $O(x^\mu \equiv 0)$.

Given two events A, B labeled by

$$x_A^\mu = (ct_A, x_A, y_A, z_A), \quad x_B^\mu = (ct_B, x_B, y_B, z_B),$$

we may then define a relative position vector of B with respect to A:

$$\Delta x^\mu = x_B^\mu - x_A^\mu = (\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3) = (c\Delta t, \Delta x, \Delta y, \Delta z).$$

Using this notation, the Lorentz transformation of the four coordinate differences Δx^μ (or their infinitesimal form dx^μ) is given, in the standard configuration, by (see also Eqs. (1.58)–(1.61) and (1.63)):

$$\begin{cases} \Delta x'^0 = \gamma(V)\left(\Delta x^0 - \frac{V}{c}\Delta x^1\right), \\ \Delta x'^1 = \gamma(V)\left(\Delta x^1 - \frac{V}{c}\Delta x^0\right), \\ \Delta x'^2 = \Delta x^2, \\ \Delta x'^3 = \Delta x^3, \end{cases} \tag{2.41}$$

which, in matrix form, can be rewritten as:

$$\begin{pmatrix} \Delta x'^0 \\ \Delta x'^1 \\ \Delta x'^2 \\ \Delta x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x^0 \\ \Delta x^1 \\ \Delta x^2 \\ \Delta x^3 \end{pmatrix}, \tag{2.42}$$

where, as usual, $\beta = V/c$. Restricting ourselves to a *standard configuration* we shall provisionally call *four-vector* any set of four quantities that, under a *standard Lorentz transformation*, undergoes the transformation (2.41) in Minkowski space. In particular, recalling Eq. (2.36), we see that the four quantities[8]

$$p^\mu = (p^0, p^1, p^2, p^3) \equiv (E/c, p_x, p_y, p_z),$$

are the components of a four-vector, the *energy-momentum* vector, which transforms by the same matrix (2.42) as (Δx^μ). Since $p^0 = E/c = m\,\gamma(v)\,c$, recalling Eq. (2.7), the *energy-momentum* vector can also be written as

$$p^\mu = m\,\frac{dx^\mu}{d\tau} = m\,\gamma(v)(c, v_x, v_y, v_z) = m\,U^\mu, \tag{2.43}$$

where U^μ, called *four-velocity*, is also a four-vector, since the rest mass m is an invariant.

Recall that, in Eqs. (1.75) and (1.76), we defined as *proper distance* in Minkowski space the Lorentz-invariant quantity

$$\Delta\ell^2 = (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2 - (\Delta x^0)^2 \equiv -c^2\,\Delta\tau^2. \tag{2.44}$$

which is the natural extension to Minkowski space of the Euclidean three-dimensional distance in Cartesian coordinates. However in the following we shall mostly use as *space-time* or *four-dimensional distance*[9] in Minkowski space the quantity $\Delta s^2 = c^2\,\Delta\tau^2 = -\Delta\ell^2$, that is the *negative* of the proper distance. This choice is dictated by the conventions we shall introduce in the following chapters when discussing the geometry of Minkowski space. Thus, for example, the square of the *four-dimensional distance or norm* of the four-vector Δx^μ is defined as

$$\|\Delta x^\mu\|^2 = (\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 = c^2\,\Delta\tau^2 = \Delta s^2. \tag{2.45}$$

Note, however, that the square of the Lorentzian norm is not *positive definite*, that is, it is not the sum of the squared components of the vector (see Eq. (2.44)) as

[8] As for Δx^μ we define $p^0 = E/c$ so that all the four components of p^μ share the same physical dimension.

[9] Alternatively also the denominations Lorentzian or Minkowskian distance are used.

the Euclidean norm $|\Delta\mathbf{x}|^2$ is. Consequently a non-vanishing four-vector can have a vanishing norm.

In analogy with the relative position four-vector, we define the norm of the energy-momentum vector as

$$\|p^\mu\|^2 = (p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = (p^0)^2 - |\mathbf{p}|^2.$$

From Eq. (2.38) it follows that this norm is precisely the (Lorentz-invariant) squared rest mass of the particle times c^2: $\|p^\mu\|^2 = m^2 c^2$. Using the notation of four-vectors, we may rewrite the results obtained so far in a more compact way.

Consider once again a collision between two particles with initial energies and momenta E_1, E_2 and $\mathbf{p}_1$, $\mathbf{p}_2$, respectively, from which two new particles are produced, with energies and momenta E_3, E_4, $\mathbf{p}_3$, $\mathbf{p}_4$. The conservation laws of energy and momentum read:

$$\begin{aligned} E_1 + E_2 &= E_3 + E_4, \\ \mathbf{p}_1 + \mathbf{p}_2 &= \mathbf{p}_3 + \mathbf{p}_4. \end{aligned} \tag{2.46}$$

If we now introduce the four-vectors p_n^μ, $n = 1, 2, 3, 4$ associated with the initial and final particles

$$p_n^\mu = \begin{pmatrix} E_n/c \\ p_{nx} \\ p_{ny} \\ p_{nz} \end{pmatrix},$$

and define the *total energy-momentum* as the sum of the corresponding four-vectors associated with the two particles before and after the process, we realize that the conservation laws of energy and momentum are equivalent to the statement that the total energy momentum four-vector is conserved. To show this we note that Eqs. (2.46) can be rewritten in a simpler and more compact form as the *conservation law of the total energy-momentum four-vector*:

$$p_{tot}^\mu = p_1^\mu + p_2^\mu = p_3^\mu + p_4^\mu. \tag{2.47}$$

Indeed the 0th component of this equation expresses the conservation of energy, while the components $\mu = 1, 2, 3$ (spatial components) express the conservation of linear momentum. Note that for each particle the norm of the energy-momentum four-vector gives the corresponding rest mass:

$$\|p_n^\mu\|^2 = \left(\frac{E_n}{c}\right)^2 - |\mathbf{p}_n|^2 = m_n^2 c^2.$$

Until now we have restricted ourselves to Lorentz transformations between frames in *standard configuration*. For the next developments it is worth generalizing our setting to Lorentz transformations with generic relative velocity vector $\mathbf{V}$, however keeping, for the time being, the three coordinate axes parallel and the origins coincident at the time $t = t' = 0$. Consider two events with relative position four-vector $\Delta x \equiv (\Delta x^\mu) = (c\,\Delta t,\ \Delta \mathbf{x})$ with respect to a frame S. We start decomposing the three-dimensional vector $\Delta \mathbf{x}$ as follows

$$\Delta \mathbf{x} = \Delta \mathbf{x}_\perp + \Delta \mathbf{x}_\parallel,$$

where $\Delta \mathbf{x}_\perp$ and $\Delta \mathbf{x}_\parallel$ denote the components of $\Delta \mathbf{x}$ orthogonal and parallel to $\mathbf{V}$, respectively. Consider now the same events described in a RF S' moving with respect to S at a velocity $\mathbf{V}$. It is easy to realize that the corresponding Lorentz transformation can be written as follows

$$\Delta \mathbf{x}' = \Delta \mathbf{x}_\perp + \gamma(V)\left(\Delta \mathbf{x}_\parallel - \mathbf{V}\Delta t\right), \tag{2.48}$$

$$\Delta t' = \gamma(V)\left(\Delta t - \frac{\Delta \mathbf{x}\cdot\mathbf{V}}{c^2}\right). \tag{2.49}$$

Indeed they leave invariant the fundamental Eq. (1.50) or, equivalently, the proper time (and thus the proper distance):

$$c^2\,\Delta t'^2 - |\Delta \mathbf{x}'|^2 = c^2\,\Delta t^2 - |\Delta \mathbf{x}|^2. \tag{2.50}$$

Writing $\Delta \mathbf{x}_\perp = \Delta \mathbf{x} - \Delta \mathbf{x}_\parallel$, $\gamma(V) \equiv \gamma$ and using the variables $\Delta x^0 = c\,\Delta t$ and $\boldsymbol{\beta} = \frac{\mathbf{V}}{c}$, Eqs. (2.48) and (2.49) become:

$$\Delta \mathbf{x}' = \Delta \mathbf{x} + (\gamma - 1)\,\Delta \mathbf{x}_\parallel - \gamma\,\boldsymbol{\beta}\,\Delta x^0, \tag{2.51}$$

$$\Delta x'^0 = \gamma\left(\Delta x^0 - \Delta \mathbf{x}\cdot\boldsymbol{\beta}\right). \tag{2.52}$$

Recalling that the four-vector $p \equiv (p^\mu) = (\frac{E}{c}, \mathbf{p})$ transforms as $x \equiv (x^\mu) = (ct, \mathbf{x})$, we also obtain

$$\mathbf{p}' = \mathbf{p} + (\gamma - 1)\,\mathbf{p}_\parallel - \gamma\,\mathbf{V}\,\frac{E}{c^2}, \tag{2.53}$$

$$E' = \gamma\,\left(E - \mathbf{p}\cdot\mathbf{V}\right), \tag{2.54}$$

and since $\mathbf{p} = m(v)\mathbf{v} = \frac{E}{c^2}\mathbf{v}$, the energy transformation (2.54) can be written as follows:

$$E' = \gamma\left(E - \frac{\mathbf{v}\cdot\mathbf{V}}{c^2}\,E\right). \tag{2.55}$$

Observing that the vector $\Delta\mathbf{x}_{\parallel}$ can also be written as $\frac{\boldsymbol{\beta}\cdot\Delta\mathbf{x}}{|\boldsymbol{\beta}|^2}\,\boldsymbol{\beta}$, the matrix form corresponding to Eqs. (2.51) and (2.52) is

$$\begin{pmatrix}\Delta x'_0\\ \Delta x'_1\\ \Delta x'_2\\ \Delta x'_3\end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta^1 & -\gamma\beta^2 & -\gamma\beta^3 \\ -\gamma\beta^1 & 1+\frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^1\beta^1 & \frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^1\beta^2 & \frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^1\beta^3 \\ -\gamma\beta^2 & \frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^2\beta^1 & 1+\frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^2\beta^2 & \frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^2\beta^3 \\ -\gamma\beta^3 & \frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^3\beta^1 & \frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^3\beta^2 & 1+\frac{(\gamma-1)}{|\boldsymbol{\beta}|^2}\beta^3\beta^3 \end{pmatrix} \begin{pmatrix}\Delta x_0\\ \Delta x_1\\ \Delta x_2\\ \Delta x_3\end{pmatrix}. \tag{2.56}$$

In the sequel we shall use the following abbreviated notation for the matrix (2.56):

$$\Lambda'^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & -\beta^j\gamma \\ -\beta^i\gamma & \delta^{ij} + (\gamma-1)\frac{\beta^i\beta^j}{|\boldsymbol{\beta}|^2} \end{pmatrix} \tag{2.57}$$

where $i, j = 1, 2, 3$ label the rows and columns of the 3×3 matrix acting on the spatial components x^1, x^2, x^3, and

$$\beta^i = \frac{v^i}{c} \Rightarrow \gamma = \frac{1}{\sqrt{1-\beta^2}},$$

where we have defined $\beta \equiv |\boldsymbol{\beta}|$. The symbol δ^{ij} is the Kronecker delta defined by the property:

$$\delta^{ij} = 1 \text{ if } i = j, \qquad \delta^{ij} = 0 \text{ if } i \neq j.$$

In Chap. 4 it will be shown that the most general Lorentz transformation $\Lambda^{\mu}{}_{\nu}$, $\mu\nu = 0, 1, 2, 3$ is obtained by multiplying the matrix $\Lambda'^{\mu}{}_{\nu}$ by a matrix $\mathbf{R} \equiv (R^{\mu}{}_{\nu})$

$$\mathbf{R} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & R^i_j \end{pmatrix} \tag{2.58}$$

where the 3×3 matrix R^i_j describes a generic rotation of the three axes (x, y, z), so that $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}'\,\mathbf{R}$. It is in terms of this general matrix that the notion of *four-vector* is defined: *A four-vector is a set of four quantities that under a general Lorentz transformation transform with the matrix* $\Lambda^{\mu}{}_{\nu}$. For example Δx^{μ} and p^{μ} are both four-vectors; indeed they have the same transformation properties under a general Lorentz transformation

$$\begin{aligned} p'^{\mu} &= \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu}\, p^{\nu}, \\ \Delta x'^{\mu} &= \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu}\, \Delta x^{\nu}. \end{aligned} \tag{2.59}$$

To simplify our notation, let us introduce the *Einstein summation convention*: Whenever in a formula a same index appears in upper and lower positions,[10] summation over that index is understood and the two indices are said to be *contracted (or dummy) indices*.[11] Using this convention, when we write for instance $\Lambda^{\mu}{}_{\nu}\, p^{\nu}$, summation over the repeated index ν will be understood, so that:

$$\Lambda^{\mu}{}_{\nu}\, p^{\nu} \equiv \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu}\, p^{\nu}.$$

Using (2.59) it is now very simple to show in a concise way that the conservation of total momentum $\mathbf{P}$ implies the conservation of the total energy E and viceversa.

Let us consider the collision of an isolated system of N particles each having a linear momentum p_n^i, $i, j = 1, 2, 3$ and let us denote by $P^i \equiv \sum_{n=1}^{N} p_n^i$ their total momentum. For each $i = 1, 2, 3$ the change ΔP^i of the total momentum $P^i \equiv \sum_{n=1}^{N} p_n^i$ occurring during the collision will be:

$$\Delta P^i \equiv \left(\sum_n p_n^i\right)_{fin} - \left(\sum_n p_n^i\right)_{in}. \tag{2.60}$$

We assume that the total momentum is conserved in a certain frame, say S, that is:

$$\Delta P^i = 0, \qquad \forall i = 1, 2, 3, \tag{2.61}$$

and carry out a Lorentz transformation to a new frame S'. Taking into account that the momentum of each particle transforms as

$$p_n^{\prime i} = \Lambda^i{}_j\, p_n^j + \Lambda^i{}_0\, p_n^0, \tag{2.62}$$

in the new frame S' the change of the total momentum is:

$$\begin{aligned}\Delta P^{\prime i} &= \left(\sum_n p_n^{\prime i}\right)_{fin} - \left(\sum_n p_n^{\prime i}\right)_{in} = \Lambda^i{}_j \left[\left(\sum_n p_n^j\right)_{fin} - \left(\sum_n p_n^j\right)_{in}\right] \\ &\quad + \Lambda^i{}_0 \left[\left(\sum_n p_n^0\right)_{fin} - \left(\sum_n p_n^0\right)_{in}\right],\end{aligned} \tag{2.63}$$

where Einstein's summation convention is used and summation over the repeated index $j = 1, 2, 3$ is understood. Since the first term in square brackets on the right

[10] So far the position of indices in vector components and matrices has been conventionally fixed. We shall give it a meaning in the next chapters.

[11] We observe that contracted indices, *being summed over*, can be denoted by arbitrary symbols, for example $\Lambda^{\mu}{}_{\nu}\, p^{\nu} \equiv \Lambda^{\mu}{}_{\rho}\, p^{\rho}$.

hand side of Eq. (2.63) is zero by hypothesis, $\Delta P^i = 0$, requiring conservation of the total momentum in S', $\Delta P'^i = 0$, implies:

$$\left(\sum_n p_n^0\right)_{fin} = \left(\sum_n p_n^0\right)_{in}, \tag{2.64}$$

that is, since

$$\sum_n p_n^0 = c \sum m_n(v) = \frac{E_{tot}}{c}, \tag{2.65}$$

the total mass, or equivalently the total energy, must be also conserved.

Viceversa, if we start assuming the conservation of energy $\sum_n p_n^0 = \sum_n E_n/c$, in S and write, for each n,

$$p_n^{0'} = \Lambda^0{}_i\, p_n^i + \Lambda^0{}_0\, p_n^0, \tag{2.66}$$

the same Lorentz transformation gives:

$$\begin{aligned}\frac{1}{c}\Delta E'_{tot.} = \left(\sum_n p_n'^0\right)_{fin} - \left(\sum_n p_n'^0\right)_{in} &= \Lambda^0{}_j \left[\left(\sum_n p_n^j\right)_{fin} - \left(\sum_n p_n^j\right)_{in}\right] \\ &+ \Lambda^0{}_0 \left[\left(\sum_n p_n^0\right)_{fin} - \left(\sum_n p_n^0\right)_{in}\right].\end{aligned} \tag{2.67}$$

The second term in square brackets on right hand side is $\Delta E_{tot}/c$ and is zero, by assumption; Being $\Lambda^0{}_j$ the three components of an arbitrary vector for generic relative motions between the two frames, each of their coefficients must vanish separately. We then conclude that the energy is conserved in S' if and only if also the total linear momentum is.

The notion of four-vector can be also used to give generalize the relativistic vector Eq. (2.15) in a four-vector notation. Recalling the invariance of the proper time interval, we may also define the *force four-vector* or *four-force* as:

$$f^\mu \equiv \frac{dP^\mu}{d\tau}. \tag{2.68}$$

To understand the content of this equation, let us consider, at a certain instant, an inertial reference frame S' moving at the same velocity as the particle. In this frame the particle will thus appear *instantaneously* at rest (the reason for not considering the rest frame of the particle, namely the frame in which the particle is *constantly at rest* is that such frame is, in general, accelerated, and thus *not inertial*). We know from our discussion of proper time that $d\tau$ coincides with the time dt' in the particle frame S' so that Eq. (2.68) becomes:

$$f'^0 = m\frac{d^2x'^0}{dt'^2} = mc\,\frac{d^2t'}{dt'^2} \equiv 0, \tag{2.69}$$

$$f'^i = m\frac{d^2x'^i}{dt'^2} \equiv F^i, \tag{2.70}$$

where we have used $dt' = d\tau$. It follows that in S' $f'^0 \equiv 0$ so that Eq. (2.70) becomes the ordinary Newtonian equation of classical mechanics. To see what happens in a generic inertial frame S, we perform a Lorentz transformation from S' to S and find:

$$f^\mu = \sum_{\nu=0}^{3} \Lambda^\mu{}_\nu\, f'^\nu = \sum_{i=1}^{3} \Lambda^\mu{}_i\, f'^i \equiv \sum_{i=1}^{3} \Lambda^\mu{}_i\, F^i. \tag{2.71}$$

The content of this equation is better understood by writing its time ($\mu = 0$) and spatial ($\mu = i$) components separately (here, for the sake of simplicity, we neglect in the Lorentz transformation the rotation part):

$$\begin{aligned} f^0 &= \Lambda^0{}_i\, F^i = \frac{\gamma}{c}\mathbf{v}\cdot\mathbf{F}, \quad f^i = \sum_{j=1}^{3} \Lambda^i{}_j\, f'^j = \sum_{j=1}^{3} \Lambda^i{}_j\, F^j \\ &= \sum_{j=1}^{3}\left(\delta^i_j + (\gamma-1)\,\frac{v^i\, v^j}{v^2}\right) F^j \equiv F^i + (\gamma-1)\,\frac{v^i}{v^2}\mathbf{v}\cdot\mathbf{F}, \end{aligned} \tag{2.72}$$

$\mathbf{v}$ being the velocity of the particle.

We observe that the expression $\mathbf{v}\cdot\mathbf{F}$ on the right hand side of the first of Eq. (2.72), is the *power* of the force $\mathbf{F}$ acting on the moving particle. The time component of Eq. (2.68) then reads

$$f^0 = \frac{dP^0}{d\tau} = \frac{\gamma}{c}\frac{dE}{dt} \equiv= \frac{\gamma}{c}\mathbf{v}\cdot\mathbf{F}, \tag{2.73}$$

and is the familiar statement that the rate of change of the energy in time equals the power of the force.

If no force is acting on the particle the equation of the motion reduces to:

$$\frac{dP^\mu}{d\tau} = 0, \tag{2.74}$$

or, using Eq. (2.43),

$$\frac{dU^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2} = 0, \tag{2.75}$$

Being $d\tau = \sqrt{1-\frac{v^2}{c^2}}\,dt$, it is easy to see that this equation implies $\mathbf{v} = \text{const}$. Thus Eq. (2.75) is the Lorentz covariant way of expressing the principle of inertia.

2.2.2 Relativistic Theories and Poincarè Transformations

We have seen that the conservation of the four-momentum and Eq. (2.68) defining the four-force are relations between four-vectors and therefore they automatically satisfy the principle of relativity being covariant under the Lorentz transformations implemented by the matrix $\mathbf{\Lambda} \equiv (\Lambda^\mu{}_\nu)$. It follows that the laws of mechanics discussed in this chapter, *excluding the treatment of the gravitational forces*, satisfy the principle of relativity, implemented in terms of general Lorentz transformations.

We may further extend the covariance of relativistic dynamics by adding transformations corresponding to constant shifts or translations

$$x'^\mu = x^\mu + b^\mu, \tag{2.76}$$

b^μ being a constant four-vector. This transformation is actually the four-dimensional transcription of time shifts and space translations already discussed for the extended Galilean transformations (1.15). However, differently from the Galilean case, there is no need in the relativistic context to add three-dimensional rotations, since, as mentioned before, they are actually part of the general Lorentz transformations implemented by the matrix $\mathbf{\Lambda}$.

It is easy to realize that the four-dimensional translations do not affect the proper time or proper distance definitions, nor the fundamental equations of the relativistic mechanics, Eqs. (2.47) and (2.68).

We conclude that relativistic dynamics is covariant under the following set of transformations

$$x'^\mu = \Lambda^\mu{}_\nu \, x^\nu + b^\mu, \tag{2.77}$$

which are referred to as *Poincaré transformations*. Both the Lorentz and Poincaré transformations will be treated in detail in Chap. 4. Furthermore in Chap. 5 it will be shown that also the Maxwell theory is covariant under the transformation (2.77), thus proving that the whole of the relativistic physics, namely relativistic dynamics and electromagnetism, is invariant under Poincaré transformations.

One could think that the invariance under translations and time shifts should not play an important role on the interpretation of a physical theory. On the contrary we shall see that such invariance *implies* the conservation of the energy and momentum in the Galilean case and of the four-momentum in the relativistic case (see Chap. 8).

2.2.3 References

For further reading see Refs. [1, 11, 12].

Springer
http://www.springer.com/978-3-319-22013-0

From Special Relativity to Feynman Diagrams
A Course in Theoretical Particle Physics for Beginners
D'Auria, R.; Trigiante, M.
2016, XVI, 601 p., Hardcover
ISBN: 978-3-319-22013-0