# Chapter 1

# Special Relativity

In the far-reaching theory of *Special Relativity* of Einstein, the homogeneity and isotropy of the 3-dimensional space are generalized to include the time dimension as well. The space-time structure embodied in the theory provides the foundation on which all branches of modern physics are formulated.[1]

## 1.1 The geometry of space-time

The basic tenet of the theory of relativity is that there is a fundamental symmetry between the three space dimensions and the time dimension, as manifested most directly in the constancy of the velocity of light in all coordinate frames. In order to formulate this theory mathematically, it is useful to introduce a set of convenient definitions and notations.

Definition 1.1 (Event): An *event*, characterised by the spatial coordinates $\{x^i; i = 1, 2, 3\}$ and the time $t$, will be denoted by $\{x^\mu; \mu = 0, 1, 2, 3\}$ where

$$x^{\mu=0} = ct, \quad x^{\mu=i} = x^i, \tag{1.1}$$

and $c$ is the velocity of light in vacuum.[2] In particle physics, the natural units are chosen, whereby $c = 1$ (by definition). The convention is that Greek indices refer to space-time in general (hence range over 0 to 3), and Roman indices refer to 3-space only (hence range over 1 to 3). We shall use the notation $\mathbf{x}$ to indicate a 3-vector.

Definition 1.2 (Coordinate Four-vector, Length of Vectors): Let $x_1^\mu$ and $x_2^\mu$ represent two events. The separation between the two events defines a *coordinate four-vector* $x_1^\mu = x_1^\mu - x_2^\mu$. The *length* $|x|$ of a 4-vector $x$ is defined by

$$|x|^2 \equiv (x^0)^2 - (\mathbf{x})^2 = t^2 - \mathbf{x} \cdot \mathbf{x}. \tag{1.2}$$

[1] Only in the theory of gravitation is it necessary to generalize the concepts of space-time beyond the special theory of relativity. The resulting theory of General Relativity is intimately related to the group of general coordinate transformations. We shall not venture into that theory in this course.

[2] In this notation, all four coordinates carry the same scale dimension - the dimension of length.

The coordinates $x^\mu$ of an event can be considered as a 4-vector if we understand it to mean the difference between the event and the event represented by the origin $(0, \mathbf{0})$. In this notation, the wave-front of a light signal sent out from the space origin at $t = 0$ will satisfy the simple equation, $|x| = 0$.

In terms of the *metric tensor* $g^{\mu\nu}$, the definition of the length of a vector $x$ can be written as

$$|x|^2 = g_{\mu\nu} x^\mu x^\nu, \tag{1.3}$$

where the implicit summations extend over all 4 components.[3]

The metric tensor for the space-time vector space is called the *Minkowski metric*:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.4}$$

In contrast to the Euclidean metric $\delta_{\mu\nu}$, the Minkowski metric is not positive definite. This fact leads to important differences in the representations of the associated symmetry groups. The principle of special relativity stipulates that basic laws of physics are invariant with respect to translations in all 4 coordinates (homogeneity of space-time) and to all homogeneous linear transformations on the space-time coordinates which leave the length of 4-vectors invariant (isotropy of space-time).

Definition 1.3 (Homogeneous Lorentz Transformation): *Homogeneous Lorentz Transformations* are continuous linear transformations $\Lambda$ on coordinate components given by

$$x^\mu \quad \to \quad x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{1.5}$$

which preserve the length of 4-vectors, i.e.

$$|x|^2 = |x'|^2 \tag{1.6}$$

Combining Eqs. 1.3-1.6, one can formulate the condition on Lorentz transformations $\Lambda$ without referring to any specific 4-vector as either

$$g_{\mu\nu} \Lambda^\mu{}_\lambda \Lambda^\nu{}_\sigma = g_{\lambda\sigma} \tag{1.7}$$

or

$$\Lambda^\mu{}_\lambda \Lambda^\nu{}_\sigma g^{\lambda\sigma} = g^{\mu\nu}, \tag{1.8}$$

where $g^{\mu\nu} = g_{\mu\nu}$. This result is an apparent generalization of rotation in 3-dimensional Euclidean space. Suppressing the indices in Eq. 1.7, we can write it (in the matrix form) as

$$\Lambda^{-1} = g \Lambda^T g^{-1}, \tag{1.9}$$

[3] This is known as the Einstein summation convention, whereby whenever an index appears twice in a product, once as a superscript, once as a subscript, the term is summed over all allowed values of that index. Since such an index does not represent any particular value, it is often called a “dummy” index or variable.

which is to be compared with $R^{-1} = R^T$ for rotation in the 3-dimensional Euclidean space.

Taking the determinant on both sides of Eq. 1.9, we obtain $(\det(\Lambda))^2 = 1$, hence $\det(\Lambda) = \pm 1$.

An example of a "large" Lorentz transformation with $\det(\Lambda) = -1$ is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.10}$$

which just flips the sign of the time coordinate, and is therefore known as time reversal:

$$t' = -t, \qquad x' = x, \qquad y' = y, \qquad z' = z. \tag{1.11}$$

Another "large" Lorentz transformation is parity, or space, inversion:

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{1.12}$$

so that

$$t' = t, \qquad x' = -x, \qquad y' = -y, \qquad z' = -z. \tag{1.13}$$

It was once thought that the laws of physics have to be invariant under these transformations, until it was shown experimentally in the 1950's that parity is violated in the weak interactions, specifically in the weak decays of the $^{60}$Co nucleus and of the $K^\pm$ mesons. Likewise, experiments in the 1960's on the decays of $K^0$ mesons showed that time-reversal is violated (at least if very general properties of quantum mechanics and special relativity are assumed).

However, all experiments up to now are consistent with invariance of the laws of physics under Lorentz transformations that are continuously connected to the identity transformation. Such transformations are known as "proper" Lorentz transformations. So, for these, we must have

$$\det(\Lambda) = \Lambda^0{}_\mu \Lambda^1{}_\nu \Lambda^2{}_\lambda \Lambda^3{}_\sigma \varepsilon^{\mu\nu\lambda\sigma} = 1, \tag{1.14}$$

where $\varepsilon^{\mu\nu\lambda\sigma}$ is the 4-dimensional totally antisymmetric unit tensor with $\varepsilon^{0123} = 1$ (the Levi-Civita tensor).[4] This condition can be rewritten as

$$\Lambda^\alpha{}_\mu \Lambda^\beta{}_\nu \Lambda^\gamma{}_\lambda \Lambda^\delta{}_\sigma \varepsilon^{\mu\nu\lambda\sigma} = \varepsilon^{\alpha\beta\gamma\delta}, \tag{1.15}$$

[4] That is,

$$\varepsilon^{\mu\nu\lambda\sigma} = \begin{cases} +1 & \text{if } \mu\nu\lambda\sigma \text{ is an even permutation of 0123,} \\ -1 & \text{if } \mu\nu\lambda\sigma \text{ is an odd permutation of 0123,} \\ 0 & \text{otherwise.} \end{cases}$$

We also note that, setting $\lambda = \sigma = 0$ in Eq.1.7, we obtain the condition

$$(\Lambda^0{}_0)^2 - \sum_i (\Lambda^i{}_0)^2 = 1. \tag{1.16}$$

This implies $(\Lambda^0{}_0)^2 \geq 1$, hence $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$. The two solutions represent disjoint regions of the real axis for $\Lambda^0{}_0$. Since $\Lambda^0{}_0 = 1$ for the identity transformation, continuity requires that all proper Lorentz transformations have

$$\Lambda^0{}_0 \geq 1. \tag{1.17}$$

Obviously, the other branch is associated with time reversal. To summarize, homogeneous proper Lorentz transformations are linear transformations of $4 \times 4$ matrices with $\Lambda^0{}_0 \geq 1$ that leave two special tensors, $g^{\mu\nu}$ and $\varepsilon^{\mu\nu\lambda\sigma}$ invariant.

A general homogeneous proper Lorentz transformation depends on 6 real parameters. This can be seen as follows: the $4 \times 4$ real matrix $\Lambda$ has 16 elements, that are subject to 10 independent constraints represented by Eq. 1.7.

Rotations in the 3 spatial dimensions are examples of Lorentz transformations in this generalized sense. They are of the form

$$R^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R^i{}_j & \\ 0 & & & \end{pmatrix}, \tag{1.18}$$

where $R^i{}_j$ denotes ordinary $3 \times 3$ rotation matrices. For example, the counterclockwse rotation by an angle $\alpha$ in the $x, y$ plane is represented by

$$R^i{}_j = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{1.19}$$

whence we have

$$\begin{aligned} t' &= t, \\ x' &= x\cos\alpha + y\sin\alpha, \\ y' &= -x\sin\alpha + y\cos\alpha, \\ z' &= z. \end{aligned} \tag{1.20}$$

Of greater interest to us are *special Lorentz transformations* which mix spatial coordinates with the time coordinate. The simplest of these is a *Lorentz boost* along a given coordinate axis, say the $x$-axis:

$$L_1{}^\mu{}_\nu = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{1.21}$$

or,

$$\begin{aligned} t' &= t\cosh\eta - x\sinh\eta, \\ x' &= -t\sinh\eta + x\cosh\eta, \\ y' &= y, \\ z' &= z. \end{aligned} \tag{1.22}$$

Physically, this corresponds to the transformation of a position vector from the unprimed frame to the primed frame, the latter moving with respect to the former along the $x$ direction at the speed $\beta = \tanh\eta$.[5] By defining

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}, \tag{1.23}$$

we have

$$\sinh\eta = \beta\gamma, \qquad \cosh\eta = \gamma. \tag{1.24}$$

Thus,

$$\begin{aligned} t' &= \gamma(t - \beta x), \\ x' &= \gamma(-\beta t + x), \\ y' &= y, \\ z' &= z. \end{aligned} \tag{1.25}$$

So, a Lorentz boost along the $x$-axis by the velocity $\beta$ can be interpreted as a "rotation" in the $t, x$ plane by the hyperbolic angle $\eta = \tanh^{-1}(\beta)$, called *rapidity*.[6]

A general Lorentz transformation can be written as the product of spatial rotations and Lorentz boosts.[7]

Definition 1.4 (Minkowski Space): The 4-dimensional space-time endowed with the Minkowski metric, Eq. 1.4, is called the *Minkowski space*. Any 4-component object $a^\mu$, transforming under Lorentz transformations as the coordinate vector in Eq. 1.5 is said to be a *four-vector* or a *Lorentz vector*.

Definition 1.5 (Scalar Product): The *scalar product* of two 4-vectors $a^\mu$ and $b^\mu$ is defined as

$$a \cdot b \equiv g_{\mu\nu} a^\mu b^\nu = a^0 b^0 - \mathbf{a} \cdot \mathbf{b}. \tag{1.26}$$

Definition 1.6 (Covariant and Contravariant Components): By convention, the ordinary components of a Lorentz vector $\{a^\mu\}$ are referred to as the *contravariant components*. An alternative way to represent the same vector is by

[5] Of course, the two frames must coincide for $\beta = 0$.

[6] Note: $0 \le |\beta| \le 1$, $1 \le \gamma$.

[7] The set of all proper Lorentz transformations $\{\Lambda\}$ satisfying the conditions of Eqs. 1.7, 1.15 and 1.17 forms the *Proper Lorentz Group* or, in short, the *Lorentz Group*. The group consists of all *special "orthogonal"* $4 \times 4$ matrices - the quotation marks here call attention to the non-Euclidean signature of the invariant metric $g^{\mu\nu}$, $(1, -1, -1, -1)$. Thus, $\Lambda$-matrices for Lorentz boosts are not unitary like the rotation matrices.

its *covariant components* $\{a_\mu\}$ defined as

$$a_\mu \equiv g_{\mu\nu} a^\nu. \tag{1.27}$$

So, $a_0 = a^0$, and $a_i = -a^i$, $i = 1, 2, 3$. With these definitions, we can simplify the definition of the scalar product, Eq. 1.26 to

$$a \cdot b = a^\mu b_\mu = a_\mu b^\mu. \tag{1.28}$$

The covariant components of a 4-vector $a$ transform under proper Lorentz transformations as

$$a_\mu \quad \rightarrow \quad a'_\mu = a_\nu (\Lambda^{-1})^\nu{}_\mu. \tag{1.29}$$

This result displays the transformation property of $a_\mu$ in the form which most explicitly indicates why $a^\mu a_\mu$ is an invariant. There is a natural covariant 4-vector, the 4-gradient $\partial_\mu$ defined as

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = (\frac{\partial}{\partial t}, \vec{\nabla}). \tag{1.30}$$

We can verify that

$$\partial_\mu \quad \rightarrow \quad \partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu{}_\mu \frac{\partial}{\partial x^\nu}. \tag{1.31}$$

With respect to an arbitrarily chosen coordinate origin, space-time is divided into three distinct regions separated by the *light-cone* which is defined by the equation

$$\tau^2 \equiv x^\mu x_\mu = t^2 - \mathbf{x} \cdot \mathbf{x} = 0. \tag{1.32}$$

The *future* consists of all points with $\tau^2 > 0$ and $x^0 > 0$. These points can be reached by the "world line" of an event at the origin. The *past* consists of all points with $\tau^2 > 0$ and $x^0 < 0$. Events at any of these points can, in principle, evolve through the origin. By a suitable Lorentz transformation, the coordinates of any point in these two regions can be transformed into the form $(t', \mathbf{0})$; hence these coordinate vectors are said to be *time-like*. The region outside the light-cone are characterized by $\tau^2 < 0$. For any given point in this region, there exists some Lorentz transformation which transforms the components of the coordinate vector into the form $(0, \mathbf{x}')$. Hence these coordinate vectors are said to be *space-like* and the entire region is called the space-like region (with respect to the origin). No world-line from the space-like region can evolve through the origin and vice versa. When $\tau^2 = 0$, the point is said to be *light-like* since only world lines of a light signal (a photon) connect such points to the origin. By analogy to the coordinates, an arbitrary 4-vector is said to be time-like, space-like, or light-like depending on whether $a^\mu a_\mu$ is less than, greater than, or equal to 0.

For two events occuring at $x^\mu$ and $x^\mu + d^\mu$, the scalar product of their 4-vector coordinate difference with itself,

$$(d\tau)^2 \equiv d^\mu d_\mu = (dt)^2 - (dx)^2 - (dy)^2 - (dz)^2 \tag{1.33}$$

is called the *proper interval (squared)* between the two events. The proper interval is, of course, independent of the choice of Lorentz frame. If two events $A$ and $B$ have $(d\tau)^2 < 0$, then the separation between them is space-like (i.e.they are not within each other's light cone). Their time-ordering is frame-dependent and, therefore, they cannot be causally connected. Note, however, that both can still be within the light cone of, and therefore, causally connected to, a third event $C$ if $C$ is either in the absolute past, or in the absolute future, of both $A$ and $B$.

For two frames with a relative Lorentz boost $\vec{\beta}$, Eq. 1.33 gives us the proper interval between the space origins at time $dt$. For convenience, let us choose the 4-coordinate origins to coincide, with $x$ and $x'$ axes oriented along the boost, so that the 3-velocity of the space origin of the frame moving along the positive $x$-axis of the other is $\dfrac{d\mathbf{x}}{dt} = (\beta, 0, 0)$. Then, dividing both sides of Eq. 1.33 by $(dt)^2$, we get

$$\frac{d\tau}{dt} = \sqrt{1-\beta^2} = \frac{1}{\gamma}, \tag{1.34}$$

or,

$$dt = \gamma d\tau. \tag{1.35}$$

This result is sometimes referred to as *time dilation.* It is as a consequence of this, that high-energy muons ($\Rightarrow$ traveling near the speed of light, as we shall soon see) created in collisions between high-energy cosmic ray particles (primarily protons) and atomic nuclei in Earth's upper atmosphere frequently traverse distances of $O(10$ km$)$ to reach the surface of earth even though the proper interval between the creation and decay of a muon is typically less than 2 $\mu$s, during which even light can travel no more than 600 m. This is possible because, while a muon may live only $\sim$2 $\mu$s in its own rest frame, to an observer on earth, the time interval between the production and decay of a cosmic-ray muon appears much longer owing to the very high speed at which the particle is moving.[8] This poses no contradiction to the observer in the muon's rest frame either. The $\sim$2 $\mu$s that he sees the muon before it decays is enough for it to travel to Earth because the thickness of the Earth's atmosphere, which is measured at, say 30 km, by a terrestrial observer, appears much less to him:

$$dx' = \frac{1}{\gamma} dx \tag{1.36}$$

This effect is sometimes referred to as *length contraction.*

In Newtonian mechanics, $t$ is an external (and universal) parameter. Therefore, $\beta \equiv \mathbf{v} \equiv \dfrac{d\mathbf{x}}{dt}$ is a 3-vector, i.e. it transforms like $\mathbf{x}$ in the 3-dimensional Euclidean space. Not so in the 4-dimensional Minkowski space, where $t$ itself is

[8]The proper interval between the production and decay of an unstable particle $X$ follows an exponential distribution: $N_X(t) = N_X(0)e^{-\frac{t}{\tau_X}}$, where $N_X(t)$ is the number of particles at time $t$. $\tau_X$ is a property, called the *lifetime*, of the particle. $\tau_\mu \approx 2.2$ $\mu$s.

a coordinate. Therefore, $v^\mu \equiv \frac{dx^\mu}{dt} = \frac{d^\mu}{d^0}$ is not a Lorentz 4-vector. However,

$$u^\mu \equiv \frac{d^\mu}{d\tau} = \gamma \frac{d^\mu}{dt} = \gamma v^\mu \tag{1.37}$$

is a Lorentz 4-vector, and is called the *relativistic velocity*. It can be easily shown that $u^\mu u_\mu = 1$.

Since $v^0 = 1$, $u^0 = \gamma$. Thus, expressed in terms of the relativistic 4-velocity $u^\mu$ (with $u^2 = u^3 = 0$), instead of the non-realtivistic 3-velocity $\vec{\beta}$, the Lorentz transformation of Eq. 1.25 reduces to a simpler, more intuitive form:

$$\begin{aligned} t' &= u^0 t - u^1 x, \\ x' &= -u^1 t + u^0 x, \\ y' &= y, \\ z' &= z, \end{aligned} \tag{1.38}$$

which is reminiscent of the Galilean transformation, except for the fact that space and time coordinates now mix in a symmetric manner.

## 1.2 Relativistic kinematics

The momentum 4-vector of a particle is

$$p^\mu = (E, \mathbf{p}), \tag{1.39}$$

where $E$ is the energy of the particle, and $\mathbf{p}$ its Euclidean 3-momentum.

In free space (i.e. in the absence of any external field interacting with the particle), the momentum is all-kinetic.[9] For a free particle with a non-zero mass $m$, the momentum is simply the product of its mass and velocity, just as in Newtonian mechanics:

$$p^\mu = m u^\mu \tag{1.40}$$

Thus,

$$p^\mu p_\mu = m^2. \tag{1.41}$$

Equation 1.41 holds for massless particles, such as photons, as well. Of course, for a given $p^\mu$, $u^\mu \to \infty$ as $m \to 0$. For $m = 0$, $u^0$ and at least one $u^i$ are undefined, but $p^\mu$ are finite. For a particle at rest, $u^0 = 1$, and we get $E = m(c^2 = 1)$, the equation that most people readily associate with Einstein.

In a closed system of particles, homogeneity of space-time ensures that the 4-momentum is conserved in any interaction (scattering, decay, annihilation):

$$\sum_i p^\mu = \sum_f p^\mu, \tag{1.42}$$

[9] When the total momentum is different from the kinetic momentum, the latter is sometimes denoted by $\pi^\mu$.

where the summation on the LHS runs over all particles in the initial state (before an ineraction) and the summation on the RHS runs over all particles in the final state (after an ineraction). Equations 1.42 are by far the most important (and often the only) ingredient in the study of particle interactions, especially from the experimental perspective.

Consider a process $ab \to cd$ From the 4-momenta of the 4 particles, we can form 10 Lorentz-invariant scalar products: $p_a^\mu p_a^\mu, p_a^\mu p_b^\mu, p_a^\mu p_c^\mu$, etc. However, these are subject to the following 8 constraints: first, onservation of 4-momentum results in the 4 equations

$$p_a^\mu + p_b^\mu = p_c^\mu + p_d^\mu, \tag{1.43}$$

and second, $p_i^\mu p_i^\mu = m_i^2$ for $i = a, b, c, d$. Thus, there must be two independent variables that describe the process. In non-relativistic mechanics, they are usually chosen to be the energy and the scattering angle. In particle physics, frame-independent quantities prove to be more convenient. Following the above arguments it is natural to define the following Lorentz scalars that are quadratic in the momenta:

$$s = (p_a + p_b)^2 = (p_c + p_d)^2, \tag{1.44}$$

$$t = (p_c - p_a)^2 = (p_d - p_b)^2, \tag{1.45}$$

$$u = (p_d - p_a)^2 = (p_c - p_b)^2. \tag{1.46}$$

These are called *Mandelstam variables.*

Clearly, the Mandelstam variables are invariant under time reversal. Also, "crossing" of processes merely results in the interchange of the Mandelstam variables. It is customary to denote the main physical process, i.e. $ab \to cd$ in this case, as the $s$ channel since $\sqrt{s}$ is the total CM energy. In the cross process $a\bar{c} \to \bar{b}d$, the CM energy would be what we now have as $t$, hence it is called the $t$ channel. Similarly, $a\bar{c} \to \bar{b}d$ would be the $u$ channel.

Any two of the three Mandelstam variables completely determines the third. A little algebra leads to the relation

$$s + t + u = m_a^2 + m_b^2 + m_c^2 + m_d^2 \tag{1.47}$$

Although only two are independent, we define all three for the sake of symmetry. Since $s, t, u$ are Lorentz scalars, they can be evaluated in any frame and used without change in any other. We shall deal mostly with symmetric colliders where the laboratory center-of-mass (CM) frame is stationary in the laboratory. By convention, we choose a coordinate system such that one of the particles is moving along the $z$-axis.[10] Thus, we have for the colliding particles

$$p_a^\mu = (E_a, 0, 0, p) \tag{1.48}$$

$$p_b^\mu = (E_b, 0, 0, -p) \tag{1.49}$$

[10]It makes good practical sense for this to lie in a horizontal plane, resulting in the choice of the perpendicular horizontal direction for $x$ and the vertical as $y$ to form a right-handed rectangular coordinate system for a particle detector.

and for the particles emerging out of the collision,

$$p_c^\mu = (E_c, \mathbf{p}') \tag{1.50}$$

$$p_d^\mu = (E_d, -\mathbf{p}') \tag{1.51}$$

Let us take the liberty to rotate our frame about the $z$-axis so that $\mathbf{p}'$ lies in the $zx$ plane:

$$\mathbf{p} = (p' \sin\theta, 0, p' \cos\theta), \tag{1.52}$$

where $\theta$ is the scattering angle in the usual 3-dimensional sense.

Then, in the CM system

$$s = (E_a + E_b)^2 = \left(\sqrt{m_a^2 + p^2} + \sqrt{m_b^2 + p^2}\right)^2, \tag{1.53}$$

which can be solved for $p$,

$$p^2 = \frac{\left(s - (m_a + m_b)^2\right)\left(s - (m_a - m_b)^2\right)}{4s}. \tag{1.54}$$

A little more algebra gives

$$E_a = \frac{s + m_a^2 - m_b^2}{2\sqrt{s}} \tag{1.55}$$

And similarly for $E_b$, $E_c$, $E_d$.

For the scattering angle, we have

$$\begin{aligned} t &= m_c^2 - m_a^2 - 2p_c^\mu p_a^\mu \\ &= m_c^2 - m_a^2 - 2E_c E_a - 2\mathbf{p}' \cdot \mathbf{p} \\ &= m_c^2 - m_a^2 - 2E_c E_a - 2p' p \cos\theta \end{aligned} \tag{1.56}$$

and

$$u = m_d^2 - m_a^2 - 2E_d E_a - 2p' p \cos\theta \tag{1.57}$$

For a process where the masses are negligible compared to the energies involved (as is most often the case in high-energy particle physics),

$$E_a = E_b = p = E_c = E_d = p' = \frac{\sqrt{s}}{2} \tag{1.58}$$

,

$$t = -\frac{s}{2}(1 - \cos\theta) \tag{1.59}$$

$$u = -\frac{s}{2}(1 + \cos\theta) \tag{1.60}$$

Like velocity, ordinary force, defined as the time derivative of momentum, is not a 4-vector. But the quantity

$$f^\mu \equiv \frac{dp^\mu}{d\tau} \tag{1.61}$$

is the force 4-vector in the Minkowski space.

Define 4 four-vectors which in a particular frame are given by the infinitesimal differentials

$$\begin{aligned}
A^\mu &= (dt, 0, 0, 0), \\
B^\mu &= (0, dx, 0, 0), \\
C^\mu &= (0, 0, dy, 0), \\
D^\mu &= (0, 0, 0, dz).
\end{aligned} \tag{1.62}$$

Then the 4-dimensional volume element

$$d^4x \equiv dx^0 dx^1 dx^2 dx^3 = A^\mu B^\nu C^\lambda D^\sigma \varepsilon_{\mu\nu\lambda\sigma} \tag{1.63}$$

is Lorentz invariant (since the RHS has no uncontracted 4-vector index). It follows that if $F(x)$ is a Lorentz scalar function of $x^\mu$, then the integral

$$I[F] = \int d^4x F(x) \tag{1.64}$$

is invariant under Lorentz transformations. Lagrangians in particle physics theories are defined in terms of such *action* integrals.

## 1.3 Maxwell's equations and the electromagnetic field

To formulate and solve Maxwell's equations in the framework of special relativity, we need to introduce two new 4-vectors: the charge-current density $J^\mu = (\rho, \mathbf{J})$,and the electromagnetic potential $A^\mu = (\phi, \mathbf{A})$.

In the non-relativistic framework, Maxwell's equations are written as

$$\begin{aligned}
\vec{\nabla} \cdot \mathbf{E} &= \rho, \\
\vec{\nabla} \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} &= \mathbf{J}, \\
\vec{\nabla} \cdot \mathbf{B} &= 0, \\
\vec{\nabla} \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0.
\end{aligned} \tag{1.65}$$

Adding the $\dfrac{\partial}{\partial t}$ of the first of the above equations to the divergence of the second, we get (since the divergence of a curl vanishes identically) the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \mathbf{J} = 0, \tag{1.66}$$

which guarantees the local conservation of electric charge. The solutions to Maxwell's equations in terms of the scalar and vector potentials are

$$\mathbf{E} = -\vec{\nabla} V - \frac{\partial \mathbf{A}}{\partial t}, \tag{1.67}$$

$$\mathbf{B} = \vec{\nabla} \times \mathbf{A}. \tag{1.68}$$

In 4-vector notation, the continuity equation Eq. 1.66 readily simplifies to

$$\partial_\mu J^\mu = 0. \tag{1.69}$$

The electric and magnetic fields can be written as components of the antisymmetric *electromagnetic field tensor*

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{1.70}$$

So, unlike in the 3-dimensional Euclidean space, the electric and magnetic fields are not vectors, but components of a tensor:

$$F_{0i} = \partial_0 A_i - \partial_i A_0 = -E_i, \tag{1.71}$$

$$F_{ij} = \partial_i A_j - \partial_j A_i = \varepsilon_{ijk} B^k, \tag{1.72}$$

where $\varepsilon_{ijk}$ is the Levi-Civita tensor in 3-dimensions.

Now the first two of Maxwell's equations become the relativistic wave equation

$$\partial_\mu F^{\mu\nu} = J^\nu, \tag{1.73}$$

or, equivalently,

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = J^\nu, \tag{1.74}$$

while the last two follow from the identity

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0. \tag{1.75}$$

The continuity equation Eq. 1.69 follows directly from Eq. 1.73: since the partial derirvatives commute and $F^{\mu\nu}$ and is antisymmtric, $\partial_\mu J^\mu = \partial_\mu \partial_\nu F^{\mu\nu} = 0$.

Note that $F^{\mu\nu}$ is explicitly invariant under a transformation

$$A^\mu \quad \to \quad A^\mu + \partial^\mu \lambda(x), \tag{1.76}$$

where $\lambda(x)$ is any scalar function of space-time. Since only the electromagnetic fields are physically manifested, any two 4-potentials that differ only by the 4-gradient of an arbitrary scalar function of space-time are equally valid for describing a physical process. This is a particular example of a general class of symmetry, called "gauge symmetry", that play a central role in the formulation of particle interactions within the standard model (electromagnetic, weak and strong interactions) and beyond.