CHAPTER 2

THE LORENTZ TRANSFORMATION

2.1 Vectors and Coordinates

In vector analysis we express relations between directed quantities e.g. the relation $\mathbf{F}=m\mathbf{a}$ between force mass and acceleration or the relation $dW=\mathbf{F}.\mathbf{dr}$ between work done, force and displacement in a way that is independent of any coordinate system. Thus if the vertices of a triangle are at **a,b,c,** the three medians of the triangle meet at $(\mathbf{a}+\mathbf{b}+\mathbf{c})/3$ and this result is independent of where we take the origin.

Alternatively we might define a Cartesian coordinate system with an origin O and three axes OX,OY,OZ, and then specify a point, such as **a**, by giving the components a_x, a_y and a_z. These components clearly depend on where we take the origin, how we choose the axes and also on the unit of length (inches? mm?). If a_x, a_y, a_z, are the components in one coordinate system, then in a new system in which the origin has been displaced by d_x, d_y, d_z, but the axes retain the same directions, the new components will be a_x-d_x, a_y-d_y, a_z-d_z. This is perhaps the simplest of all coordinate transformations and we could express it as

$$x\rightarrow x'= x-d_x, \quad y\rightarrow y'= y-d_y, \quad z\rightarrow z'= z-d_z \qquad (2.1)$$

Another kind of transformation would be one in which the axes were rotated leaving the origin unchanged. In this case the new coordinates would be related to the old by a linear transformation of the general form

$$x'= ax+by+cz, \quad y'= dx+ey+fz, \quad z'= gx+hy+iz \qquad (2.2)$$

where the constant coefficients a,b,c,d etc. depend on the axis and angle of rotation. Thus, for a rotation of the axes through an angle ϕ about the z axis, we have $a = e = \cos\phi$, $b = -d = \sin\phi$, $i = 1$ with all the other coefficients zero. A further type of transformation occurs when the new system moves relative to the old system with a constant velocity **u** having components u_x, u_y, u_z relative to the original coordinate system. In this case

$$x'= x-u_xt, \quad y'= y-u_yt, \quad z'=z-u_zt. \qquad (2.3)$$

This is the classical transformation of Galilean relativity.

It leaves the **form** of all the **classical** dynamic laws unaltered. Thus if a particle of mass m has a velocity $\mathbf{v}$ with components $v_x, v_y, v_z,$ in the first coordinate system and is subject to a force $\mathbf{F}$ so that $m dv_x/dt=F_x$ etc, then in the new system $F'_x=F_x$ and $m dv'_x/dt=F'_x$ etc, Of course this classical transformation would change the velocity of light c to $c-u$ and so could not be compatible with electromagnetism.

2.2 The Lorentz Transformation

Like the Galilean transformation this relates the coordinates of a single event in one inertial coordinate system to the coordinates of the same event in a second inertial coordinate system moving with a uniform velocity relative to the first system, when both systems use the same time standards and length standards. Within each system the clocks are synchronised but now we do not make any assumption that if r is the distance between two points in one system then r', the distance between the same two points referred to the second system, is equal to r, nor that if t is the time between two events in the first system then t' the time between the two events referred to the second system is also equal to t. This, as we shall see, allows us to find a unique transformation that makes the velocity of light the same in all inertial coordinate systems.

Because the only defined direction is that of the relative velocity $\mathbf{u}$ of the two frames of reference we can choose the axes so that OX and OX' coincide in direction with this velocity, and we can also choose the origin of time in both systems to occur when the two origins O and O' coincide. We also choose the OY' and OZ' axes so that, at this instant, they coincide with OY and OZ.

A set of coordinates x, y, z, t given, in the original system S, to a particular event must, at the very least, determine a unique set of coordinates x',y',z',t' in the new system S' and *vice versa.* The **only** transformation with this property of reciprocal uniqueness that does **not** transform **any** neighbouring points in F to points **infinitely** far apart in F' is a **linear** transformation. Thus we must write $x'=Ax+By+Cz+Dt, y'=Ex+Fy+Gz+Ht, z'=Ix+Jy+Kz+Lt, t'=Mx+Ny+Pz+Qt$. [In the theory of optical instruments we can also consider

collinear transformations $x_i' = \Sigma_j\{a_{ij}x_j+b_i\}/\Sigma_k\{f_kx_k+g\}$ where i,j and k run over 1,2,3 corresponding to x,y and z, and sums over j and k are implied. Here it does not matter if a point where the denominator vanishes is imaged at infinity, as this is just the property of the focus (see Appendix A8).

The successive positions of a particle starting at O and moving along the X axis with a velocity u define a series of events for which $x=ut, y=0, z=0$ in the unprimed system, but the particle is at rest at $x'=0, y'=0, z'=0$ in the primed system. It follows that since x' depends linearly on x and t, it can only do so as $x' = g(x-ut)$ where g may depend on u. Furthermore, since the equations are all linear, should y' and z' depend on t they can only do so with t in the combination $x-ut$. But the line of relative motion is the **only** defined direction in the system and since there are no other preferred directions neither y' nor z' can depend on x. They cannot therefore depend on t, similarly nor can t' depend on either y or z. Because y' does not depend on t an event on the z (or z') axis with $y=y'=0$ at $t=0$ stays on the z' axis and so y' cannot depend on z. Similarly z' cannot depend on y. Finally t' must be related to t and x by $t'=ht+kx$.

A rotation through 180° about the y axis changes x to $-x$, u to $-u$ and leaves t and t' unchanged, thus $h(-u) = h(u)$, $g(-u) = g(u)$ and $k(-u) = -k(u)$. The inverse transformation from the primed to the unprimed coordinates is obviously obtained by changing u to $-u$, for the origin 0 has a velocity $-u$ along the 0'X' axis. Apart from the trivial relation $y'=y$, $z'=z$, we have

$$x' = g(u)[x-ut], \qquad t' = h(u)t + k(u)x, \qquad (2.6a)$$

$$x = g(u)[x'+ut'], \qquad t = h(u)t' - k(u)x'. \qquad (2.6b)$$

If Eq.(2.6a) is substituted in Eq.(2.6b) the result is $x = [g^2+ugk]x+u[h-g]t$, $t = [h^2+ukg]t+[hk-gk]x$. These two equations must be identities in x and t and so we have $h=g$, and $k=[1-g^2]/ug$ leading to

$$x' = g[x-ut], \qquad t' = g[t+ x(1-g^2)/ug^2], \qquad (2.7a)$$

$$x = g[x'+ut'], \qquad t = g[t'-x'(1-g^2)/ug^2]. \qquad (2.7b)$$

These, together with the simple relations, $y'=y$, $z'=z$, are the most general transformation laws (given our choice of the axes and origins) between the coordinates of an event in two different **inertial** coordinate systems with a relative velocity u. They have arisen solely because we have assumed (i) that space and time are homogeneous, i.e. there are no preferred positions or instants, (ii) that space is also isotropic without any preferred directions and (iii) that if Newton's first law holds in one coordinate system it also holds in any other system in uniform relative motion. Apart from u they depend on one further parameter g which may be a function of u, though not, of course, of x,y,z or t. If we wish to retain the intuitive notion that $t'=t$, then we must put $g=1$ and this reduces Eqs.(2.7a,b) to the Galilean transformation of classical mechanics i.e.

$$x' = x - ut, \quad y' = y, \quad z' = z\ , \quad t' = t. \tag{2.8}$$

So far, however, the function $g(u)$ is arbitrary, except that it must be an even function of u and reduce to unity as u tends to zero.

We can now impose a further condition, the invariance of the velocity of light c. Let $x=vt$, then $x'=g(u)(v-u)t$ and $t'=g(u)[1+v(1-g^2)/ug^2]t$, giving $x'=[v-u]t'/[1+v(1-g^2)/ug^2]$. We now require that if $v=c$, then $v'\equiv x'/t'=c$. This leads to $c[1+c(1-g^2)/ug^2]=c-u$ giving the relation $g^2=1/(1-u^2/c^2)$. Since g must tend to +1 as u tends to zero we finally have

$$\gamma(u)\equiv g(u) = (1-u^2/c^2)^{-\frac{1}{2}}\ , \tag{2.9}$$

where we have replaced g by the traditional Greek letter γ to give a formula with which we shall soon become very familiar.

The complete Lorentz transformation therefore becomes

$$x'= \gamma(u)[x-ut] = (1-u^2/c^2)^{-\frac{1}{2}}[x-ut], \tag{2.10a}$$

$$y' = y, \qquad z' = z, \tag{2.10b}$$

$$t'= \gamma(u)[t-ux/c^2] = (1-u^2/c^2)^{-\frac{1}{2}}\ [t-ux/c^2], \tag{2.10c}$$

and the inverse transformation is

$$x = \gamma(u)[x'+ ut'], \tag{2.11a}$$

$$y = y', \quad z = z', \tag{2.11b}$$

$$t = \gamma(u)[t'+ux'/c^2]. \tag{2.11c}$$

These two sets of equations relate the coordinates and times (x,y,z,t) and (x',y',z',t') given to the **same** identical single event in two coordinate systems which are defined in two different inertial frames of reference, such that the origin of the primed frame moves along the OX axis of the unprimed frame with a constant velocity u, and the origins and axes of the two frames coincide at $t=t'=0$. They describe the *only* transformation that incorporates the isotropy of space (there are no preferred directions) and the uniformity of space and time (no preferred locations or instants), that preserves Newton's first law and, finally, that leaves the velocity of light c as a universal constant, the same in all inertial frames of reference.

2.3 Simple Kinematic Results

Because c, initially introduced as the velocity of light, has now appeared in equations describing space-time relations it will appear in other parts of physics having nothing to do with electromagnetism. In particular, as we shall see in a later chapter, if we wish to retain the law of conservation of linear momentum when speeds comparable with c are involved, the usual classical relation $\mathbf{p}=m\mathbf{v}$ between momentum $\mathbf{p}$, mass m and velocity $\mathbf{v}$ has to be replaced by

$$\mathbf{p} = m\mathbf{v}/(1-v^2/c^2)^{\frac{1}{2}} = m\mathbf{v}\gamma(v),$$

and this has profound repercussions throughout physics.

One way of looking at this relation between momentum and velocity is to suppose that the mass m of a particle of velocity v is related to its **rest mass** m_o by $m = m_o\gamma(v)$, and Kaufmann (1902), three years before Einstein's paper and a year before Lorentz published his transformation, had (in a version of J.J.Thomson's experiments) observed just this variation with velocity of the mass-to-charge ratio of an electron, but had ascribed it to an electromagnetic effect!

We look first at some purely kinematic consequences of the Lorentz transformation. Obviously the most surprising aspect of the transformation is that t and t', the times ascribed to the same event, are different in the two coordinate systems. This, alone, might be enough to make us have doubts about the relevance of the transformation to real physics, were it not for the way that nature provides a simple, accurate, direct and immediately comprehensible experimental demonstration of the time transformation.

Muons, particles first detected in cosmic rays, decay when they are at rest (or only moving slowly), about 2 microseconds after their formation through the decay of a pion. It is, however, also possible to observe the decay of muons with velocities approaching c. Now suppose that a muon is formed at rest in a frame F' with a velocity u relative to the laboratory frame of reference F in which the coordinates of an event are x,y,z,t (henceforth we use "coordinates" to include time as well as space). We take it, for convenience, to be formed at $t'=t=0$, and to decay at $t'=\tau$ in its own rest frame F'. Thus τ is its lifetime at rest. It also decays at the same place x' (in its rest frame) at which it was formed and therefore we might as well take $x'=0$. In the laboratory frame F it is formed at $x=0$ and decays, according to Eq. (2.11a), at $x=\gamma(u)ut'=\gamma(u)u\tau$ and, according to Eq. (2.11c), at $t=\gamma(u)\tau=\tau/(1-u^2/c^2)^{\frac{1}{2}}$. This time t can be considerably greater than τ. For example if the muons are created with a velocity 0.99c the observed life time will be some 7 times longer than the natural lifetime τ. Experiments with muons trapped in a storage ring have verified the relation

$$t = \tau/(1 - u^2/c^2)^{\frac{1}{2}} \tag{2.12}$$

up to values of t/τ greater than 100. This effect, **"time dilation"**, is what Poincaré meant by saying that moving clocks must run slow.

In the laboratory frame F, between its formation and its decay, the muon moves a distance $ut = u\tau/(1 - u^2/c^2)^{\frac{1}{2}}$ which, if u is comparable to c, can be much greater than τc. In fact muons formed in the upper atmosphere at heights up to 10 km reach the earth's surface. Even travelling with the velocity of light this would take 30 microseconds, much longer than the natural lifetime of the muon. Of course this result is also one of the staples of science fiction for it allows voyagers with a natural life-time of only three score years and ten to travel to galaxies many, many thousands of light years away (if, that is, they could be accelerated to very nearly the velocity of light!).

A second simple kinematic result is the formula for combining velocities. The trajectory of a particle travelling along the O'X' axis of the frame F' with a velocity v' is given by $x'=v't'$. Its description in the laboratory frame F is $x = \gamma(u)[x'+ut'] = \gamma(u)[v'+u]t'$ and $t = \gamma(u)[t'+ux'/c^2] = \gamma(u)[1+uv'/c^2]t'$ so that in the laboratory frame the velocity is

$$v = x/t = [v'+ u]/[1 + uv'/c^2]. \tag{2.13}$$

We note that if either u or v' is equal to c, then $v = c$.

Combining any velocity with c always gives c.

We can use Eq.(2.13) to calculate $\gamma(v)$ in terms of u and v'. We begin with $1-v^2/c^2 = [(1+uv'/c^2)-(u+v')^2/c^2]/[1-uv'/c^2]^2$ which simplifies to give

$$\gamma(v) = \gamma(u)\gamma(v')[1 + uv'/c^2]. \qquad (2.14)$$

This is itself a useful result with many applications but for the moment we note that, whatever their signs, as long as u and v' are numerically less than c, the right hand side of Eq.(2.14) is positive, finite and real, thus $\gamma(v)$ is also positive, finite and real so that v must be numerically less than c.

Combining two velocities less than c can never yield a velocity greater than, or even equal to c.

Because times and coordinates must be expressed by real numbers, the Lorentz transformation can only be a proper physical relation if $\gamma(u)=1/(1-u^2/c^2)^{1/2}$ is real and therefore if $u\leq c$. It is impossible to construct meaningful relations between frames of reference whose relative velocity exceeds c. On the other hand if a body has a velocity u relative to another body, we could certainly give a meaning to the two frames in which one or the other body is at rest. We must conclude that when $u>c$ there can be no relation between events associated with both the bodies. This can be expressed by saying that c is the greatest velocity that can appear in any equation expressing an observable relation between any two bodies. Thus not only is c the invariant velocity of light but it is also the greatest velocity with which **signals of any kind** can propagate. Any theory that ignores this leads to nonsensical paradoxes in dealing with causality. This does not mean that velocities greater than c cannot occur. As a trivial example if a pair of scissors is closed the point where the blades intersect moves forward much faster than the blades move together, and there is nothing in relativity to say that the point of intersection cannot move faster than c. As Trester (1989) points out the distinction is between the propagation of **effects** and of **signals**. It is also worth remarking that although in most cases waves propagate signals at their group velocity, this is not a universal rule. The subtle distinction between group velocity and signal velocity is discussed in detail by Brillouin (1960).

Whereas time dilation is readily observed experimentally the corresponding spatial effect, called the Lorentz contraction, is not. In general, objects large and solid enough to have a well defined length at rest are too massive to be accelerated to anywhere near the velocity of light. Nevertheless an appreciation of the effect is often helpful in understanding the content of a relativistic calculation. Consider a rod which, when at rest (in its rest frame), has a length L' and which is moving along the 0X axis parallel to its length with a velocity u. It is illuminated by a single flash of light at $t=0$ when the trailing end of the rod is at $x = 0$ and the leading end is at $x = L$. In the rod's rest frame this is at $x'=L'$, but $x'=\gamma(u)[x-ut]$ and, since $t=0$, this gives $L'=\gamma(u)L$ or

$$L = L'(1 - u^2/c^2)^{\frac{1}{2}} , \tag{2.15}$$

and the length L in the laboratory F frame is **less** than the length L' in the rod's rest frame F'. Although both ends of the rod are seen in the frame F at the same time, they are seen at different times in the frame F' moving with the rod. Suppose instead that the rod were to be illuminated all at the same time $t' = 0$ in the frame F'. Then the leading edge of the rod would appear at $x = \gamma(u)[x'+ ut'] = \gamma(u)L'$ giving

$$L = L'/(1 - u^2/c^2)^{\frac{1}{2}} \tag{2.16}$$

and the rod would appear **longer**, but now of course the two ends are seen in the laboratory frame F at different times $t=0$ and $t=\gamma(u)uL'/c^2=uL/c^2$, and between these two times the rod has moved a distance u^2L/c^2. As we remarked earlier, the Lorentz contraction of a macroscopic body is not a directly observable experimental effect: the energy needed to accelerate it to anything like the velocity of light is prohibitive.

The Lorentz transformation only relates the coordinates to be ascribed to events in two different frames of reference in relative motion, it does not tell us directly how an object specified in one frame, for example as a solid cube, might appear in the other frame. Discussions of the appearance of an object moving with nearly the velocity of light have very little relevance to the world of observable physical phenomena.