Chapter 5

Momentum and Energy

Having seen that the nineteenth century theory of Maxwell is, in spite of initial appearances, a Lorentz invariant theory, we may now address the problem of fixing the seventeenth century theory of Newton so that it too becomes a Lorentz invariant theory. How is this to be done?

Our objective is to learn something about the behavior of material bodies subjected to forces. The only guide we have is electromagnetic theory, so let's consider electromagnetic forces. For this purpose we shall consider a body of finite extent composed of charges and currents, and we shall subject this body to a given electromagnetic field. According to Maxwell's theory the force density exerted upon the body is given by

$$\vec{F} = \rho\vec{E} + \frac{1}{c}\vec{\jmath}\times\vec{H} \ .$$

The actual force would be this expression integrated over the volume of the body. If we then integrate the force over a time interval, the result will be the *impulse* or change $\Delta\vec{P}$ in the momentum of the body during the time interval in question. That is,

$$\Delta\vec{P} = \int\int\int\int [\rho\vec{E} + \frac{1}{c}\vec{\jmath}\times\vec{H}]\ dx\ dy\ dz\ dt \ .$$

On the other hand, if we wish to compute the change ΔE in the energy of the body during the time interval in question, we shall have to integrate the *power* over that time interval. The power is, however, the rate at which energy is transferred to the body as a result of Joule heating; i.e., the integral of $\vec{\jmath}\cdot\vec{E}$ over the volume of the body. (Remember that for a simple material $\vec{E} = \sigma\vec{\jmath}$, where σ is the conductivity.) Thus, the change in energy is given by

$$\Delta E = \int\int\int\int [\vec{\jmath}\cdot\vec{E}]\ dx\ dy\ dz\ dt \ .$$

Can we infer anything concerning how momentum and energy must transform under Lorentz transformations from our knowledge of how the electromagnetic objects transform? To answer this we must study the quantities $\rho\vec{E}+\frac{1}{c}\vec{\jmath}\times\vec{H}$ and $\vec{\jmath}\cdot\vec{E}$. Both of these quantities are linear in electromagnetic fields and linear in charge and current densities. This suggests trying to form a vector field from the skew-symmetrical tensor field

$$\mathbf{F}=H_x\mathbf{e}_2\wedge\mathbf{e}_3+H_y\mathbf{e}_3\wedge\mathbf{e}_1+H_z\mathbf{e}_1\wedge\mathbf{e}_2-E_x\mathbf{e}_1\wedge\mathbf{e}_4-E_y\mathbf{e}_2\wedge\mathbf{e}_4-E_z\mathbf{e}_3\wedge\mathbf{e}_4$$

and the vector field

$$\mathbf{J}=j_x\mathbf{e}_1+j_y\mathbf{e}_2+j_z\mathbf{e}_3+c\rho\mathbf{e}_4\ .$$

We have already extended the notion of inner product from the scalar product of two vectors to the scalar product of two bivectors. Now we shall extend the notion of inner product to the inner product of a bivector and a vector. The definition is simply

$$(\mathbf{A}\wedge\mathbf{B})\cdot\mathbf{C}:=\mathbf{B}(\mathbf{A}\cdot\mathbf{C})-\mathbf{A}(\mathbf{B}\cdot\mathbf{C})\ .$$

With this definition of the inner product, what is the value of the inner product $\mathbf{F}\cdot\mathbf{J}$? Substituting the expressions for $\mathbf{F}$ and $\mathbf{J}$, we soon find that

$$\frac{1}{c}\mathbf{F}\cdot\mathbf{J}=(\rho\vec{E}+\frac{1}{c}\vec{\jmath}\times\vec{H})\cdot\vec{\mathbf{e}}+\frac{1}{c}(\vec{\jmath}\cdot\vec{E})\mathbf{e}_4\ .$$

In other words, the components of the force density $\vec{F}$ are just the first three components of a four-vector, the fourth component of which is $1/c$ times the density of Joule heating.

Since the volume element $dx\ dy\ dz\ dt$ is Lorentz invariant, it follows that the momentum change $\Delta\vec{P}$, and $1/c$ times the energy change ΔE, are components of a four-vector, say $\Delta\mathbf{P}$, where

$$\mathbf{P}:=P_x\mathbf{e}_1+P_y\mathbf{e}_2+P_z\mathbf{e}_3+\frac{1}{c}E\mathbf{e}_4\ .$$

This is the hint we need in order to replace Newton's theory by a Lorentz invariant theory. It tells us that momentum and energy divided by c must transform as components of a four-vector $\mathbf{P}$. That is certainly not true for Newton's theory, where the energy E and the momentum $\vec{P}$ of a particle are related by $E=P^2/2M$, where M is the mass of the particle. What is the *correct* relation between E and $\vec{P}$?

P is an invariant geometrical object, a four-vector. Its "length" $\mathbf{P}\cdot\mathbf{P}$ is an invariant, i.e., it has the same value in all inertial frames. If you select the frame of reference in which the particle is (instantaneously) at rest, in this frame $\vec{P}=\vec{0}$ and $E=E_0$, to which we shall refer as the *rest energy* of the particle.

Using the fact that $\mathbf{P}\cdot\mathbf{P}$ is an invariant, and is equal to $-E_0^2/c^2$, we conclude that in *any* inertial frame,

$$P^2-\frac{1}{c^2}E^2=-\frac{1}{c^2}E_0^2\ .$$

Thus, we obtain the important result that

$$E^2=E_0^2+c^2P^2\ .$$

Now, how can we reconcile this energy-momentum relation with the apparently successful formula $E=P^2/2M$ of Newton? Remember, that Newton was concerned with slowly moving particles. Suppose the rest energy of a particle E_0 were not zero. Suppose, in fact, $c^2P^2\ll E_0^2$. Then we would be justified in approximating Einstein's formula as follows:

$$E=\sqrt{E_0^2+c^2P^2}=E_0+c^2P^2/2E_0+\cdots\ ,$$

where higher order terms have been neglected. The term which is quadratic in P resembles Newton's expression $P^2/2M$. If we identify the rest energy as

$$E_0=Mc^2\ ,$$

then we will have

$$E=\sqrt{(Mc^2)^2+c^2P^2}=Mc^2+\frac{P^2}{2M}+\cdots\ .$$

Until the twentieth century, the rest energy term remained completely unobserved. Only with the advent of the study of particle reactions in general, and atomic energy in particular, was the importance of the Mc^2 term appreciated.

In classical mechanics there is a simple relation $\vec{P}=M\vec{v}$ between momentum and velocity. In relativity we need a four-vector relationship, one involving the four-momentum **P**. Is there a four-vector velocity **V** such that $\mathbf{P}=M\mathbf{V}$?

One immediately thinks of taking a limit of the four-vector

$$\Delta\mathbf{X}:=\Delta x\mathbf{e}_1+\Delta y\mathbf{e}_2+\Delta z\mathbf{e}_3+c\Delta t\mathbf{e}_4$$

divided by the elapsed time, as the elapsed time shrinks to zero. However, whose clock is to be used in determining the elapsed time? The only way we shall get a four-vector is if we use a Lorentz invariant notion of elapsed time. We must, therefore, use the *proper time* $\Delta\tau$, as it would be measured using a clock attached to the particle. Hence, we shall write

$$\mathbf{V} := \frac{d\mathbf{X}}{d\tau} \ .$$

This can be expressed in terms of ordinary velocity $\vec{v}$ very easily, if we recall the time-dilation effect,

$$\Delta\tau = \Delta t \sqrt{1 - v^2/c^2}.$$

Thus, we have

$$\mathbf{V} = \frac{\vec{v} \cdot \vec{\mathbf{e}} + c\mathbf{e}_4}{\sqrt{1 - v^2/c^2}} \ ,$$

from which it follows that

$$\begin{aligned} \vec{P} &= \frac{M\vec{v}}{\sqrt{1 - v^2/c^2}} \\ E &= \frac{Mc^2}{\sqrt{1 - v^2/c^2}} \ . \end{aligned}$$

In the low velocity limit these formulas reduce to

$$\begin{aligned} \vec{P} &= M\vec{v} + \cdots \ , \\ E &= Mc^2 + \frac{1}{2}Mv^2 + \cdots \ . \end{aligned}$$

To get some idea of the magnitude of the terms, an electron has a rest energy $Mc^2 = 0.5$ MeV. An *electron volt* is the energy acquired when an electron is accelerated by a potential difference of one volt. When electrons come out of the Stanford linear accelerator they have energies around 500 MeV, i.e., about 1000 times their rest energy. At these speeds one has to use the correct relativistic relation between E and v. The classical approximation is no good.

Compton Scattering

As a practical application of Einstein's energy-momentum relationship, let us consider the phenomenon of *Compton scattering.* Here an X-ray is scattered by a free electron, initially at rest. The unquantized Maxwell theory is completely inadequate to describe the observations, for if the classical theory were correct, one would expect no shift in the wave length of the scattered light, regardless of the angle of scattering. The electron would simply oscillate in a simple harmonic fashion in response to the oscillating $\vec{E}$ field of the incident electromagnetic wave, and the oscillating electron would emit spherical waves (dipole radiation) of the same frequency. In fact, what Compton observed was that there is a shift in the wave length of the scattered light, a shift given by the empirical formula

$$\Delta\lambda = 2\lambda_0 \sin^2(\theta/2) ,$$

where θ is the angle of scattering and the constant λ_0 is equal to 0.0242 Ångstroms. (1 Ångstrom = 10^{-8} cm.) To understand this result we must use Einstein's corpuscular view picture of light coupled with his special theory of relativity.

According to relativity theory, the total four-momentum is conserved. If $\mathbf{P}_1$ and $\mathbf{P}_2$ are the four-momenta of the initial electron and the recoil electron, respectively, and $\mathbf{k}_1$ and $\mathbf{k}_2$ are the four-momenta of the incident and outgoing photon, respectively, then

$$\mathbf{P}_1 + \mathbf{k}_1 = \mathbf{P}_2 + \mathbf{k}_2 .$$

It follows that

$$(\mathbf{P}_1 - \mathbf{P}_2)^2 = (\mathbf{k}_1 - \mathbf{k}_2)^2 ,$$

where we are computing the scalar products of four-vectors with themselves, i.e., Lorentz invariant quantities.

Expanding this result, we have

$$\mathbf{P}_1^2 + \mathbf{P}_2^2 - 2\mathbf{P}_1 \cdot \mathbf{P}_2 = \mathbf{k}_1^2 + \mathbf{k}_2^2 - 2\mathbf{k}_1 \cdot \mathbf{k}_2 . \tag{5.1}$$

However, according to the relativity theory,

$$\mathbf{P}_1^2 = \mathbf{P}_2^2 = -(mc)^2 \text{ and } \mathbf{k}_1^2 = \mathbf{k}_2^2 = 0 ,$$

where m is the mass of the electron, and the photon mass is assumed to vanish. If, furthermore, we denote the energy of the initial electron by E_1 and that of the recoil electron by E_2, the energy of the incident photon by $\mathcal{E}_1$ and that of the scattered photon by $\mathcal{E}_2$, then we have

$$\begin{aligned} \mathbf{P}_1 \cdot \mathbf{P}_2 &= \vec{P}_1 \cdot \vec{P}_2 - E_1 E_2/c^2 , \\ \mathbf{k}_1 \cdot \mathbf{k}_2 &= \vec{k}_1 \cdot \vec{k}_2 - \mathcal{E}_1 \mathcal{E}_2/c^2 . \end{aligned}$$

We now assume that the initial electron is at rest, so that $\vec{P}_1 = \vec{0}$ and $E_1 = mc^2$. Furthermore, we introduce the scattering angle θ, the angle between $\vec{k}_1$ and $\vec{k}_2$. Then

$$\vec{k}_1 \cdot \vec{k}_2 = |\vec{k}_1||\vec{k}_2| \cos\theta = \mathcal{E}_1\mathcal{E}_2/c^2 \ .$$

Substituting all these expressions into Eq. (5.1), we obtain

$$2m(E_2 - mc^2) = \frac{\mathcal{E}_1\mathcal{E}_2}{c^2} \sin^2(\frac{\theta}{2}) \ .$$

However, by conservation of energy,

$$mc^2 + \mathcal{E}_1 = E_2 + \mathcal{E}_2 \ ,$$

so one can easily show that

$$\frac{1}{\mathcal{E}_2} - \frac{1}{\mathcal{E}_1} = \frac{2}{mc^2} \sin^2(\frac{\theta}{2}) \ .$$

Following Einstein's suggestion and setting $\mathcal{E}_1 = h\nu_1$ and $\mathcal{E}_2 = h\nu_2$, we obtain

$$\lambda_2 - \lambda_1 = 2(\frac{h}{mc}) \sin^2(\frac{\theta}{2}) \ ,$$

where we have used the relation $\nu\lambda = c$ between frequency and wave length.

This is precisely the Compton scattering formula, with the empirical constant λ_0 identified with $\lambda_c := h/mc$, a combination of fundamental constants that is called the Compton wave length and whose value agrees very well with the experimental value $\lambda_0 = 0.0242$ Ångstroms. Indeed, I find it amazing that only one new fundamental constant h had to be introduced not only to describe Compton scattering, but the spectrum of hydrogen, the photoelectric effect and the Planck blackbody radiation formula as well. Nature is sometimes extremely kind to us. But let us not become too cocky because of our past accomplishments!