Relativity Notes

David Murdock, TTU

September 27, 2011

Chapter 1

Preliminaries

We do a quick review of the background material in classical physics which you should bring to these notes (or that you should be learning in other courses) so that this stuff will be meaningful.

1.1 Newton's Laws

Physics largely involves the measurement of motion and Newtonian physics stresses the importance of position, velocity and acceleration:

$$\mathbf{r}(t) \qquad \mathbf{v}(t) = \frac{d\mathbf{r}}{dt} \qquad \mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2}$$

Newton gave three laws of motion for objects small enough so that their size doesn't matter ("particles") which in their simplistic first–year form are given as:

• In the absence of forces, an object will move with constant velocity.

• When forces act on an object of mass m, the acceleration of the object is given by

$$\mathbf{F}_{\mathrm{net}} = m\mathbf{a}$$

where $\mathbf{F}_{\mathrm{net}}$ is found by taking the vector sum of the individual forces which act on the mass:

$$\mathbf{F}_{\mathrm{net}} = \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 + \cdots$$

• A force on an object always originates from *some other object* in the universe and if object B exerts a force $\mathbf{F}_{AB}$ on object A and object A exerts a force $\mathbf{F}_{BA}$ on object B then

$$\mathbf{F}_{AB} = -\mathbf{F}_{BA}$$

In addition Newton gave us the law for the only fundamental force he knew, that of gravity:

• Two point masses m_1 and m_2 which are separated by a distance r exert an attractive force on one another of magnitude

$$F_{\rm grav} = G\frac{m_1 m_2}{r^2} \qquad \text{where} \qquad G = 6.67 \times 10^{-11}\,\frac{\rm N{\cdot}m^2}{\rm kg^2}$$

and by "attractive" we mean the force is directed along the line joining the two masses and toward the other mass.

1.2 Electromagnetism

It took a while longer for people to learn about another fundamental force between particles, the force of electromagnetism. We introduce the entities known as the electric and magnetic fields, $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$ and with these, if a particle has charge q and moves with velocity $\mathbf{v}$ in the presence of these fields, then the force on this particle is

$$\mathbf{F} = q[\mathbf{E} + \mathbf{v} \times \mathbf{B}]$$

The electric and magnetic fields arise from a charge density $\rho(\mathbf{r},t)$ and a current density $\mathbf{J}(\mathbf{r},t)$ and the fields and sources are related by the Maxwell equations (written in "MKS units"),

$$\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{1}{\epsilon_0}\rho \\
\nabla \cdot \mathbf{B} &= 0 \\
\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0 \\
\nabla \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} &= \frac{1}{\mu_0}\mathbf{J}
\end{aligned}$$

Here, ϵ_0 and μ_0 are related to the speed of light c by $c = 1/\sqrt{\epsilon_0\mu_0}$ and one can demonstrate that electric charge and current obey an equation of continuity:

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{1.1}$$

As we'll see in more detail later, there is a wonderful symmetry among these equations which can only be appreciated with Einstein's theory of relativity.

With these laws of physics in hand (and learning the empirical rules rules for other less fundamental forces in the world) one can go forth and solve many problems of motion and understand the electric and magnetic nature of matter. Up to a point, they can even be applied to the motion of atoms to help understand the thermal properties of matter, although to go very far in that direction we soon need to use quantum physics. Nevertheless, the laws of classical physics do a fine job in describing the macroscopic world. At the end of the 19th century many physicists thought they had all the laws of nature in hand and only detailed calculations remained.

But not all physicists were so convinced. The smartest ones of the time realized that there were serious problems with this picture: One was that the laws did not apply very well to the workings of atoms, whose existence was then becoming widely accepted. The question of reference frames, especially in regard to electromagnetism, was also very troubling to some physicists, as will be discussed in the next chapter. Einstein's great achievement was to put the question of reference frames front and center in physics and show how this consideration determines what laws of physics are *possible*.

Chapter 2

Reference Frames and the Laws of Physics

2.1 Introduction

In these notes I'll assume you've heard *something* about Einstein's theory of special relativity even if all you did with it was to derive some astonishing things about poles moving through barns at high speeds. These notes will *not* discuss poles and barns. Not for a while, at any rate.

Einstein's Theory of Relativity is perhaps best described as a *research program* for physics. That is, it is not about any *particular* theory. It gives a scheme for seeking out the correct (mathematical) laws of nature by giving restrictions on their possible forms, a restriction based on reference frames. With these restrictions in place the equations of physics can be successfully applied to particles which are moving at speeds close to that of light. And when implemented in either classical and quantum physics, special relativity has implications which are far–reaching and often counter-intuitive.

Even if the work *you* do in physics does not directly use the principles of special relativity, it is important for physics as a whole.

2.2 What Is a Reference Frame?

A reference frame is a platform (real or hypothetical) on which we make measurements of space and time for particles in motion and thereby do physics; specifically we measure *events*, with their space and time coordinates: (x, y, z, t).

Physics is *necessarily* done in various reference frames so it is necessary to have a way of comparing the measurements made in the different frames. We also need to know what the laws of physics *are* in the different frames; The issue can't be avoided if only because we live in a reference frame (the rotating earth) which almost certainly is not the "correct" reference frame of the universe— even if there is such a thing.

Throughout the first part of these notes, for simplicity and to standardize our choices we will discuss a frame $\bar{\mathcal{S}}$ with axes $\bar{x}$, $\bar{y}$ and $\bar{z}$ pointing in the same directions as those of the frame $\mathcal{S}$. (I will use a bar to denote different reference frames to avoid the possible

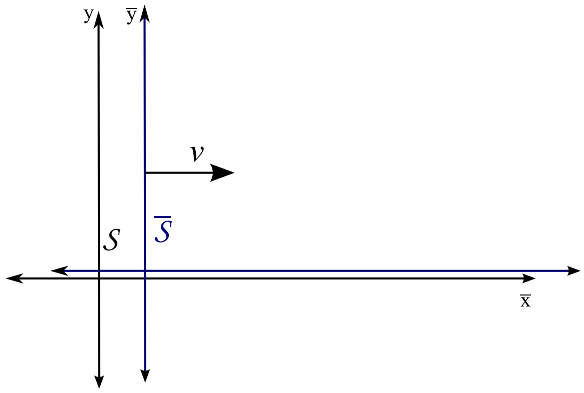

Figure 2.1: Reference frames.

confusion that a prime might designate a derivative.) Frame $\bar{\mathcal{S}}$ moves in the $+x$ direction with respect to frame $\mathcal{S}$ at a *constant* speed v, and the origins of the two systems coincide at the zero of time as measured in each system. We can later generalize to frames moving in other directions or even frames where the axes are rotated with respect to one another. (The case where v is not constant leads to further interesting results but that's way too much to consider for right now.)

In each reference frame, measurements of distance and time are made and the physicists who live in each frame try to formulate and test the laws of physics.

2.3 Galilean Transformation

The idea that reference frames moving with respect to one another can't be distinguished by testing the fundamental laws of physics is not new; it even occured to Galileo, but there is a lot of work to do to implement this principle: What *are* the fundamental laws of physics and how is the condition of invariance expressed mathematically? One particular choice for the application of this principle is that of *Galilean relativity*.

Concerning our two frames $\mathcal{S}$ and $\bar{\mathcal{S}}$, common sense dictates that time t will have the same value in both frames (time is universal, isn't it?) and that the lengths measured along the x axis at any given time will simply add; this leads to the simple **Galilean transformation** between length and time measurements made in the two frames.

$$\bar{x} = x - vt \qquad \bar{y} = y \qquad \bar{z} = z \qquad \bar{t} = t \tag{2.1}$$

Though this is simple, for completeness, we note the reverse of this transformation,

$$x = \bar{x} + vt \qquad y = \bar{y} \qquad z = \bar{z} \qquad t = \bar{t} \tag{2.2}$$

which *had* to follow from 2.1 by replacing v by $-v$, but it also follows by simple algebra!

2.4 Invariance of Newton's 2nd Law

We examine the form that Newton's 2nd law takes in the different frames. Suppose a particle has a trajectory given by $\mathbf{r}(t)$ in frame $\mathcal{S}$. Then in frame $\bar{\mathcal{S}}$ its trajectory is given by

$$\bar{\mathbf{r}}(\bar{t}) = \mathbf{r}(\bar{t}) - \mathbf{v}\bar{t} = \mathbf{r}(t) - \mathbf{v}t$$

where $\mathbf{v} = v\hat{\mathbf{i}}$. The velocity of the particle in frame $\mathcal{S}$ we denote by $\mathbf{u}$ (saving $\mathbf{v}$ for the relative velocity of the frames):

$$\mathbf{u}(t) = \frac{d}{dt}\mathbf{r}(t)$$

while in frame $\bar{\mathcal{S}}$ it is

$$\begin{aligned} \bar{\mathbf{u}}(\bar{t}) &= \frac{d}{d\bar{t}}\bar{\mathbf{r}}(\bar{t}) = \frac{d}{dt}\left[\mathbf{r}(\bar{t}) - \mathbf{v}\bar{t}\right] \\ &= \mathbf{u}(t) - \mathbf{v} \end{aligned}$$

so the two frames disagree on the velocity of a particle (in exactly the way we expect), but they will agree on its acceleration, since with

$$\mathbf{a} = \frac{d}{dt}\mathbf{u}(t)$$

we have

$$\begin{aligned} \bar{\mathbf{a}}(\bar{t}) &= \frac{d}{d\bar{t}}\bar{\mathbf{u}}(\bar{t}) = \frac{d}{dt}(\mathbf{u}(t) - \mathbf{v}) \\ &= \mathbf{a}(t) \end{aligned}$$

Again, we are assuming that $\mathbf{v}$, the relative velocity of the frames is *constant*.

We will suppose that if in frame $\mathcal{S}$ the force at point $\mathbf{r}$ and time t is $\mathbf{F}(\mathbf{r}, t)$ then in frame $\bar{\mathcal{S}}$ the observers will agree that the same force $\mathbf{F}$ acts on the particle. (It will be a different *mathematical function* of the time and space coordinates but the same *vector quantity*.) This assumption will be questioned in the pages ahead!

Then the equation $\mathbf{F} = m\mathbf{a}$ will hold in both frames because we have shown that $\bar{\mathbf{a}} = \mathbf{a}$ and we have supposed that $\bar{\mathbf{F}} = \mathbf{F}$. Newton's 2^{nd} law is said to be invariant between inertial reference frames: If it holds in *any one frame* then it will hold in a frame which moves at constant velocity with respect to that one. Even the claim that Newtons' law "holds" is subtle; true forces must always be due to *other particles* in the universe and will behave according the Newton's third law.

With this understanding we see that Newton's 1st law has more much content than one would guess from the way it is presented in elementary physics courses. Correctly stated, it is *not* just a special case of the second law (if it were, why would we need both?) It can be taken as the statement that there is *some* reference frame where a lack of forces gives a lack of acceleration, and then Newton's 2nd law *applies to that frame* but also any other frame moving at constant velocity.

2.5 The Maxwell Equations

Another set of equations of classical physics that one might take to be fundamental are the Maxwell equations for the electromagnetic field. The Maxwell equations do *not* transform in a simple way under the Galilean transformation — that is, once we change the coordinates over to those of the $\bar{\mathcal{S}}$ system there's no simple way to make the result look like the original Maxwell equations. But as we'll see, they do have a very interesting property when subjected to the Lorentz transform).

2.6 But What's "Fundamental", Anyway?

Sometimes there is a preferred reference frame for a particular phenomenon which does not arise from anything truly *fundamental* in the laws of physics. For example, to analyze the motion of sound waves in air, a unique and simple reference frame is provided by the rest frame of the air but there's nothing fundamental about this phenomenon.

Similarly, at one time it was conceivable that there a preferred frame for electromagnetic phenomena. In the late 19th century, electric and magnetic fields were thought to be something like sound waves. The fields were imagined as some kind of (abstract) deformation in some kind of (abstract) medium, known as the **ether**. This idea was not crazy for its time; it was a natural extension of the laws of physics as then understood. Unfortunately it was wrong. But in such a situation we would not expect the Maxwell equations to have the same form in all reference frames, because there is only one right reference frame for them, i.e. the rest frame of the ether. Maybe the frame of the laboratories on earth isn't *too* much different from that frame so that we have to look hard to see the effects of our motion. In particular, the speed of a light wave, which arises in a simple way from the Maxwell equations, might well have different values in different reference frames just as the speed of sound does as observers move with respect to the air. It was very hard for physicists of the late 19th century to get away from the idea that the ether— or some *mechanical basis for electromagnetism*— was necessary. It was hard for physicists then *and today* to accept the idea that the speed of light is the same in all reference frames. But nature does not care what we like or what we can easily wrap our minds around.

2.7 Lorentz Transformation

The relativistic transformation of space and time coordinates was named after H. A. Lorentz, who recognized the utility of these equations in the analysis of electromagnetism but didn't grasp their true meaning; it took Einstein to do that.

We first define the all–important quantity γ,

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{2.3}$$

where v is the relative speed of two reference frames and where we will assume that $v < c$ as there is no evidence to suggest that any particle (and thus any proper frame of reference) can move faster than the speed of light with respect to any other frame. So we note that γ is always larger than 1 and increases as $v \to c$.

We will sometimes use $\beta \equiv \frac{v}{c}$ so that $\gamma = \frac{1}{\sqrt{1-\beta^2}}$.

$$\bar{x} = \gamma(x - vt) \qquad \bar{y} = y \qquad \bar{z} = z \qquad \bar{t} = \gamma\left(t - \frac{v}{c^2}x\right) \tag{2.4}$$

with the reverse transformation coming from solving these for (x, y, z, t) or more simply changing the sign of v:

$$x = \gamma(\bar{x} + v\bar{t}) \qquad y = \bar{y} \qquad z = \bar{z} \qquad t = \gamma\left(\bar{t} + \frac{v}{c^2}\bar{x}\right) \tag{2.5}$$

It will be noticed immediately that the time t is no longer a universal quantity but now is mixed in with the coordinate (x) along the direction of the relative motion. The concept of simultaneity now requires much more care and there are other famous (nonintuitive) effects involving space and time which we'll look at later!

2.8 Addition of Velocities

The notation we will use here is that a particle's velocity in a particular frame is $\mathbf{u}$ to carefully distinguish it from the velocity of the frame itself ($\mathbf{v}$, that is, the velocity of $\bar{\mathcal{S}}$ with respect to $\mathcal{S}$). Later on we can combine the two ideas by having a reference frame move along with a particle but we should keep them separate at first! In the two frames the components of the velocity are defined as

$$u_x \equiv \frac{dx}{dt} \qquad u_y \equiv \frac{dy}{dt} \qquad u_z \equiv \frac{dz}{dt} \qquad \bar{u}_x \equiv \frac{d\bar{x}}{d\bar{t}} \qquad \bar{u}_y \equiv \frac{d\bar{y}}{d\bar{t}} \qquad \bar{u}_z \equiv \frac{d\bar{z}}{d\bar{t}}$$

and where the fact that $\bar{t} \neq t$ is where all the fun begins!

An easy way derive the relation between $\bar{\mathbf{u}}$ and $\mathbf{u}$ is to consider that moving particle as being at two points closely separated in space in frame $\mathcal{S}$ by (dx, dy, dz) and in time by dt. The components of the particle's velocity in frame $\mathcal{S}$ are

$$u_x = \frac{dx}{dt} \qquad u_y = \frac{dy}{dt} \qquad u_z = \frac{dz}{dt}$$

The two events are separated in frame $\bar{\mathcal{S}}$ by spatial separations $(d\bar{x}, d\bar{y}, d\bar{z})$ and a time separation $d\bar{t}$. These are related to the unbarred intervals by

$$d\bar{x} = \gamma(dx - vd\bar{t}) \qquad d\bar{y} = dy \qquad d\bar{z} = dz \qquad d\bar{t} = \gamma(dt - \frac{v}{c^2}dx)$$

We take a look outside to spot any math teachers who don't like us dividing differential and then find the x component of velocity in frame $\bar{\mathcal{S}}$:

$$\bar{u}_x = \frac{d\bar{x}}{d\bar{t}} = \frac{\gamma(dx - vdt)}{\gamma(dt - \frac{v}{c^2}dx)} = \frac{\frac{dx}{dt} - v}{(1 - \frac{v}{c^2}\frac{dx}{dt})} = \frac{u_x - v}{(1 - \frac{u_x v}{c^2})}$$

Likewise we get $\bar{u}_y$ and $\bar{u}_z$:

$$\bar{u}_y = \frac{d\bar{y}}{d\bar{t}} = \frac{dy}{\gamma(dx - vdt)} = \frac{u_y}{\gamma(1 - \frac{u_x v}{c^2})} \qquad \bar{u}_z = \frac{d\bar{z}}{d\bar{t}} = \frac{dz}{\gamma(dx - vdt)} = \frac{u_z}{\gamma(1 - \frac{u_x v}{c^2})}$$

Collecting these results, we have

$$\bar{u}_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}} \qquad \bar{u}_y = \frac{u_y}{\gamma(1 - \frac{u_x v}{c^2})} \qquad \bar{u}_z = \frac{u_z}{\gamma(1 - \frac{u_x v}{c^2})} \tag{2.6}$$

Velocities do *not* add in a simple way! There are multiplicative factors which even enter into the components of velocity *perpendicular to the relative motion*. We recognize that the simple answer of the Galilean transformation (that is, $\bar{u}_x = u_x - v$) is altered and the other components are changed as well from the fact that the time variable also transforms.

It follows that if light (or any other particle) has a speed c in one frame, it has speed c in *all* frames. To see this, suppose the velocity of a photon in frame $\mathcal{S}$ is $u_x = c$ with $u_y = 0$ and $u_z = 0$. Then 2.5 gives

$$\bar{u}_x = \frac{c - v}{1 - \frac{cv}{c^2}} = \frac{c(1 - \frac{v}{c})}{(1 - \frac{v}{c})} = c$$

The math is a little more complicated if in frame $\mathcal{S}$ the photon can be moving in some general direction with speed $u = c$ such that

$$u_x^2 + u_y^2 + u_z^2 = c^2$$

Then 2.5 gives the speed in frame $\mathcal{S}$ as

$$\begin{aligned} \bar{u}^2 &= \bar{u}_x^2 + \bar{u}_y^2 + \bar{u}_z^2 \\ &= \frac{(u_x - v)^2}{\left(1 - \frac{u_x v}{c^2}\right)} + \frac{u_y^2}{\gamma^2\left(1 - \frac{u_x v}{c^2}\right)^2} + \frac{u_z^2}{\gamma^2\left(1 - \frac{u_x v}{c^2}\right)^2} \\ &= \frac{\gamma^2(u_x - v)^2 + u_y^2 + u_z^2}{\gamma^2\left(1 - \frac{u_x v}{c^2}\right)^2} \end{aligned}$$

This a mess, but the algebra of relativity is frequently messy and you need to get used to it! Using the fact that $1/\gamma^2 = 1 - \frac{v^2}{c^2}$, we can combine all terms as

$$\bar{u}^2 = \frac{(u_x - v)^2 + (u_y^2 + u_z^2)(1 - \frac{v^2}{c^2})}{\left(1 - \frac{u_x v}{c^2}\right)^2}$$

Now use $u_y^2 + u_z^2 = c^2 - u_x^2$ and this gives

$$\bar{u}^2 = \frac{(u_x - v)^2 + (c^2 - u_x^2)(1 - \frac{v^2}{c^2})}{\left(1 - \frac{u_x v}{c^2}\right)^2} = \frac{-2u_x v + c^2 + u_x^2 v^2/c^2}{\left(1 - \frac{u_x v}{c^2}\right)^2}$$

where we've cancelled a couple terms in the numerator. If we further rearrange the terms we get

$$\bar{u}^2 = \frac{c^2\left(1 - \frac{2u_x v}{c^2} + \frac{u_x^2 v^2}{c^4}\right)}{1 - \frac{2u_x v}{c^2} + \frac{u_x^2 v^2}{c^4}} = c^2$$

so we again find that a speed c in frame $\bar{\mathcal{S}}$ gives a speed c in frame $\bar{\mathcal{S}}$.

2.9 Accelerations, Newton's 2nd Law

With the Lorentz transformation accelerations of particles as measured in the different frames are *very* different. With the definitions

$$a_x = \frac{dv_x}{dt} \qquad a_y = \frac{dv_y}{dt} \qquad a_z = \frac{dv_z}{dt} \qquad\qquad \bar{a}_x = \frac{d\bar{u}_x}{d\bar{t}} \qquad \bar{a}_y = \frac{d\bar{u}_y}{d\bar{t}} \qquad \bar{a}_z = \frac{d\bar{u}_z}{d\bar{t}}$$

we get a set of rather complicated relations which are too awful for anyone to quote. I just did the math myself and found

$$\bar{a}_x = \frac{d\bar{u}_x}{d\bar{t}} = \frac{a_x}{\gamma^3\left(1 - u_x v/c^2\right)^3} = a_x \frac{(1 - v^2/c^2)^{3/2}}{\left(1 - u_x v/c^2\right)^3}$$

which approaches a_x in the limit of small v but otherwise is a mess.

Accelerations do not play the central role in relativistic dynamics that they do in Newtonian dynamics. But to be clear about this, even though the frames have a constant velocity of *relative motion* we are here discussing the change in time of the velocity of a particle as measured within each of those frames.

If anything *like* Newton's 2nd law can be shown to hold (and that ought to be true because after all it *does* work when the speeds are small), then it will require a lot more in the ways of definitions and tricks. We will do this later on.

2.10 The Maxwell Equations

We now look at the transformation of the Maxwell equations under the Lorentz transformation of Eq. 4.1. We will see that they *do* keep the same form, a fact which impressed Einstein enough to convince him that Eq. 4.1 was right and Eq. 2.1 was wrong. The derivation is a bit messy but it is worth doing at this stage of the discussion because it illustrates some important points about the transformation of a law of physics. Later we will re-do the derivation with more streamlined mathematics.

The Maxwell equations (in tedious but ubiquitous MKS units, but using $\mu_0 = \frac{1}{c^2\epsilon_0}$ for simplicity) are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{2.7}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.8}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.9}$$

$$\nabla \times \mathbf{B} = \frac{\mathbf{J}}{c^2\epsilon_0} + \frac{\partial \mathbf{E}}{\partial t} \tag{2.10}$$

Here, the vectors $\mathbf{E}$ and $\mathbf{B}$ are functions of (x, y, z, t). These equations are assumed to hold in frame $\mathcal{S}$.

How do we transform these to another reference frame? It's a subtle question which I'd like to discuss before diving into the math. Certainly the derivatives in 2.6 through 2.9 are with respect to (x, y, z, t) need to be changed into the $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ system and we must now

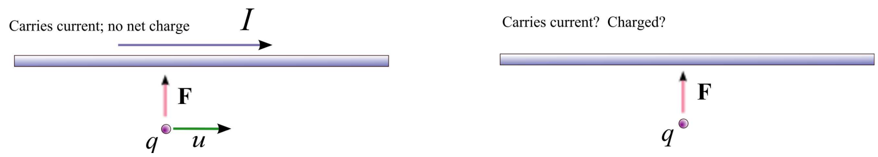

Figure 2.2: Two reference frames in which a charge can "view" a current–carrying or charged (?) wire. In the first, the charge is in motion. In the second, the charge is at rest.

work (carefully) with partial derivatives. But when we get done with this, will we expect the electric field $\mathbf{E}$ of frame $\mathcal{S}$ playing the *exact* same role in frame $\bar{\mathcal{S}}$?

As it turns out, this is too much to expect and there are simple physics reasons why. Consider a lab situation where a charge q moves at speed u parallel to a current–carrying (but electrically neutral) wire as shown in Fig. 2.10. For a positive charge q, the charge will experience a force given by $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$; it is attracted to the wire. But if we consider a reference frame which moves along with the particle, we would guess that even in this frame we have a force on the charge toward the wire (possibly not the same as in the first picture). But it can't be a magnetic force since as we see it, the charge has no velocity. If the laws of electricity are still valid, what kind of force could it be? It would seem that it could only be an *electric* force but this could only arise if in the new frame the wire is *not* electrically neutral (though it may still carry a current). So between the two frames it would seem that we won't agree on what the electromagnetic fields are, and we don't agree on the charge density of the wire.

We have to expect that in frame $\bar{\mathcal{S}}$ the $\mathbf{E}$ and $\mathbf{B}$ fields will be *some mixture* of the fields in $\mathcal{S}$, and preferably not a very complicated one.

This derivation will be long and thankless, but let's do it!

First off we need a dictionary for changing partial deriviates with respect to (x, y, z) to ones with respect to $(\bar{x}, \bar{y}, \bar{z})$. Using some basic multi-variable calculus and the Lorentz transformation in Eq. 4.1, this is:

$$
\begin{aligned}
\frac{\partial}{\partial x} &= \frac{\partial \bar{x}}{\partial x}\frac{\partial}{\partial \bar{x}} + \frac{\partial \bar{t}}{\partial x}\frac{\partial}{\partial \bar{t}} = \gamma\frac{\partial}{\partial \bar{x}} + \gamma v\frac{\partial}{\partial \bar{t}} \\
\frac{\partial}{\partial y} &= \frac{\partial}{\partial \bar{y}} \\
\frac{\partial}{\partial z} &= \frac{\partial}{\partial \bar{z}} \\
\frac{\partial}{\partial t} &= \frac{\partial \bar{x}}{\partial t}\frac{\partial}{\partial \bar{x}} + \frac{\partial \bar{t}}{\partial t}\frac{\partial}{\partial \bar{t}} = \gamma\frac{\partial}{\partial \bar{t}} + \frac{\gamma v}{c^2}\frac{\partial}{\partial \bar{x}}
\end{aligned}
$$

When you get done with all of the algebra (included in an appendix), we find the following new versions of the $\mathcal{S}$ Maxwell equations (written in the same order as the old ones):

$$\begin{aligned}
\frac{\partial \bar{E}_x}{\partial \bar{x}} + \frac{\partial}{\partial \bar{y}}[\gamma E_y - \gamma v B_z] + \frac{\partial}{\partial \bar{z}}[\gamma E_z + \gamma v B_y] &= \frac{1}{\epsilon_0}(\gamma\rho - \frac{\gamma v}{c^2}J_x) \\
\frac{\partial \bar{B}_x}{\partial \bar{x}} + \frac{\partial}{\partial \bar{y}}[\gamma B_y + \frac{\gamma v}{c^2}E_z] + \frac{\partial}{\partial \bar{z}}(\gamma B_z - \frac{\gamma v}{c^2}E_y) &= 0
\end{aligned}$$

$$\begin{aligned}
\left(\frac{\partial}{\partial \bar{y}}[\gamma E_z + \gamma v B_y] - \frac{\partial}{\partial \bar{z}}[\gamma E_y - \gamma v B_z]\right)\hat{\mathbf{i}} &+ \left(\frac{\partial E_x}{\partial \bar{z}} - \frac{\partial}{\partial \bar{x}}[\gamma E_z + \gamma v B_y]\right)\hat{\mathbf{j}} \\
+\left(\frac{\partial}{\partial \bar{x}}[\gamma E_y - \gamma v B_z] - \frac{\partial E_x}{\partial \bar{y}}\right)\hat{\mathbf{k}} &+ \frac{\partial}{\partial \bar{t}}\left(B_x\hat{\mathbf{i}} + [\gamma B_y + \frac{\gamma v}{c^2}E_z]\hat{\mathbf{j}} + [\gamma B_z - \frac{\gamma v}{c^2}E_y]\hat{\mathbf{k}}\right) = 0 \\
\left(\frac{\partial}{\partial \bar{y}}[\gamma B_z - \frac{\gamma v}{c^2}E_y] - \frac{\partial}{\partial \bar{z}}[\gamma B_y + \frac{\gamma v}{c^2}E_z]\right)\hat{\mathbf{i}} &+ \left(\frac{\partial B_x}{\partial \bar{z}} - \frac{\partial}{\partial \bar{x}}[\gamma B_z - \frac{\gamma v}{c^2}E_y]\right)\hat{\mathbf{j}} \\
+\left(\frac{\partial}{\partial \bar{x}}[\gamma B_y + \frac{\gamma v}{c^2}E_z] - \frac{\partial B_x}{\partial \bar{y}}\right)\hat{\mathbf{k}} &- \frac{1}{c^2}\frac{\partial}{\partial \bar{t}}\left(E_x\hat{\mathbf{i}} + [\gamma E_y - \gamma v B_z]\hat{\mathbf{j}} + [\gamma E_z + \gamma v B_y]\hat{\mathbf{k}}\right) \\
&= \frac{1}{c^2\epsilon_0}([\gamma J_x - \gamma v\rho]\hat{\mathbf{i}} + J_y\hat{\mathbf{j}} + J_k\hat{\mathbf{k}})
\end{aligned}$$

Ouch! This is a mess!

Or is it?

Compare these equations with the old ones. We notice the combinations of E's and B's that appear all through these equations and we find that if we define

$$\bar{E}_x \equiv E_x \qquad \bar{E}_y \equiv \gamma E_y - \gamma v B_z \qquad \bar{E}_z = \gamma E_z + \gamma v B_y \tag{2.11}$$

$$\bar{B}_x \equiv B_x \qquad \bar{B}_y \equiv \gamma B_y + \frac{\gamma v}{c^2}E_z \qquad \bar{B}_z = \gamma B_z - \frac{\gamma v}{c^2}E_y \tag{2.12}$$

$$\bar{\rho} \equiv \gamma\rho - \gamma v J_x \qquad \bar{J}_x \equiv \gamma J_x - \gamma v\rho \qquad \bar{J}_y = J_y \qquad \bar{J}_z = J_z \tag{2.13}$$

then the previous equations are not so bad:

$$\begin{aligned}
\frac{\partial \bar{E}_x}{\partial \bar{x}} + \frac{\partial \bar{E}_y}{\partial \bar{y}} + \frac{\partial \bar{E}_z}{\partial \bar{z}} &= \bar{\nabla}\cdot\mathbf{E} = \frac{1}{\epsilon_0}\bar{\rho} \\
\frac{\partial \bar{B}_x}{\partial \bar{x}} + \frac{\partial \bar{B}_y}{\partial \bar{y}} + \frac{\partial \bar{B}_z}{\partial \bar{z}} &= \bar{\nabla}\cdot\mathbf{B} = 0 \\
\left(\frac{\partial \bar{E}_z}{\partial \bar{y}} - \frac{\partial \bar{E}_y}{\partial \bar{z}}\right) &+ \left(\frac{\partial \bar{E}_x}{\partial \bar{z}} - \frac{\partial \bar{E}_z}{\partial \bar{x}}\right) + \left(\frac{\partial \bar{E}_y}{\partial \bar{x}} - \frac{\partial \bar{E}_x}{\partial \bar{y}}\right) + \frac{\partial \mathbf{B}}{\partial \bar{t}} = 0 \\
\text{or} \qquad \bar{\nabla}\times\mathbf{E} + \frac{\partial \mathbf{B}}{\partial \bar{t}} &= 0 \\
\left(\frac{\partial \bar{B}_z}{\partial \bar{y}} - \frac{\partial \bar{B}_y}{\partial \bar{z}}\right) &+ \left(\frac{\partial \bar{B}_x}{\partial \bar{z}} - \frac{\partial \bar{B}_z}{\partial \bar{x}}\right) + \left(\frac{\partial \bar{B}_y}{\partial \bar{x}} - \frac{\partial \bar{B}_x}{\partial \bar{y}}\right) - \frac{\partial \mathbf{E}}{\partial \bar{t}} = \frac{1}{c^2\epsilon_0}\bar{\mathbf{J}} \\
\text{or} \qquad \bar{\nabla}\times\mathbf{B} - \frac{\partial \mathbf{E}}{\partial \bar{t}} &= \frac{1}{c^2\epsilon_0}\bar{\mathbf{J}}
\end{aligned}$$

In fact they now look *exactly* like the Maxwell equations. What if... they *are* the Maxwell equations? Then it would mean that the electric and magnetic fields in frame $\bar{\mathcal{S}}$ can be found from those in $\mathcal{S}$ by the transformations given in 2.10.

What emerges from this lengthy calculation is surprising. We start with the space and time coordinates jumbled together and when we sort them out to look like Maxwell equations we end up with *fields* that are jumbled together, but following an interesting pattern! The pattern has a *resemblance* to the Lorentz transformation itself, and the transformation of the source terms looks *exactly* like the Lorentz transformation.

2.11 Rescuing the Laws of Motion

We can rescue Newton's law ($\mathbf{F} = m\mathbf{a}$) by recasting it in a form which does have a nice transformation property between reference frames. It's a bit too much for the current discussion but the basic plan is this:

We note that in Newton's mechanics one can write $\mathbf{F} = m\mathbf{a}$ or $\mathbf{F} = \frac{d\mathbf{p}}{dt}$; they are equivalent because the masses of *particles* in Newtonian mechanics do not change[1]. In relativity we will redefine the momentum of a particle of mass m, velocity $\mathbf{u}$ as

$$\mathbf{p} \equiv \frac{m\mathbf{u}}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{2.14}$$

and then we will use

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

and now this is *not* equivalent to $\mathbf{F} = m\mathbf{a}$.

As we'll see, we also have conservation laws in relativistic mechanics; in collisions of two particles, momentum and energy are both conserved but only when we use 2.13 for the momentum (of a massive particle) and use another formula for the energy will we have conserved quantities.

2.12 So What's the Big Deal?

So what does it really matter how we change the coordinates in one frame into those in another frame?

As we've seen, the insistence that the laws of physics keep the same mathematical form in all frames leads to a hard decision: With our two choices for the transformation of coordinates, the basic choice is that Newton is right and Maxwell is wrong or Newton is wrong and Maxwell is right. (Well, perhaps not "wrong" but "needs to be fixed".)

Still, if we have the proper laws of physics why worry about jumping between reference frames?

A good example is provided by the decay of the muon; the muon is an elementary particle which on average *in the reference frame where it is at rest* decays in $2.2 \times 10^{-6}\,\mathrm{s}$, producing an electron a two neutrinos. We can pose the question: If muons are produced such that they are traveling at a speed of $\frac{4}{5}c$, how far (on average) do we expect them to travel? The naive answer multiplies the given lifetime by the speed $\frac{4}{5}c$ to get:

$$d = v\tau = \left(\frac{4}{5}c\right)(2.2 \times 10^{-6}\,\mathrm{s}) = 5.3 \times 10^{2}\,\mathrm{m}$$

[1]Authors of some basic textbooks say that they do. No. They don't.

But this is wrong; the given lifetime was relevant for the reference frame in which the muon is at rest. In a frame which moves with respect to the muon (i.e. the lab), as we'll discuss later, this time interval is measured as

$$\tau' = \gamma\tau = \tfrac{5}{3}(2.2 \times 10^{-6}\,\mathrm{s}) = 3.7 \times 10^{-6}\,\mathrm{s}$$

so that the muons will travel a distance

$$d = v\tau' = (\tfrac{4}{5}c)(\tfrac{5}{3}\tau) = 8.8 \times 10^{2}\,\mathrm{m}$$

The person who doesn't like (or understand) relativity will be troubled by this result. We might imagine the following conversation between him and someone more used to its counter-intuitive predictions:

"But why should the muon live longer if it is moving? What is it about the muon's decay that has changed because of its motion?"

"I don't have to answer that question. It's sufficient to know about what happens in one frame and to use the Lorentz transformation rule."

"But what's the *mechanism* behind the time lengthening? Surely we have to know something about the physics of muon decay to say that."

"Nope. All we need to know is that *all* the fundamental processes of physics (muon decay being one of them) work the same way in each frame, and time and space are transformed by the Lorentz relations. Learning the *detailed theory* of muon decay will be nice, but it's not necessary for finding its lifetime in the lab, and that theory— whatever it is— will have to have the right transformation properties."

"This is frustrating."

Indeed, many students of physics do find it frustrating but Nature is the way she is. She doesn't care if we're frustrated. And initially many physicists regarded Einstein's theory the same way our skeptic does: as a cheap way of ducking the important questions in physics. When we start to learn what relativity really says, perhaps we can have a little more sympathy for those physicists.

Chapter 3

Space–Time

3.1 Introduction

It was clear that when we did calculations for the transformation between reference frames some patterns were being followed in the way that the components of space and time or the components of the E and B fields mixed together. This pattern arises from what we may think of as a generalized rotation of the coordinates and the mathematics behind it all is both aesthetically pleasing and useful for calculations.

And we'll explore it in this chapter.

3.2 Vectors in 2 Dimensions; Rotations.

You know about vectors in 2 dimension: The basic operations of addition, finding components, the dot product of two vectors and all that. I don't need to review it.

But what I do want to review is the mathematics that goes with a rotation of the coordinate system, as shown in Fig. 3.2. We will use the same "bar" notation for the different coordinate systems that we used in Chapter 1: One coordinate system will be called $\mathcal{S}$ and the other $\bar{\mathcal{S}}$; $\bar{\mathcal{S}}$ is rotated with respect to $\mathcal{S}$ by a counterclockwise angle ϕ.

One can easily show that the coordinates of a given point are related by

$$x = x\cos\phi + y\sin\phi \qquad\qquad y = -x\sin\phi + y\cos\phi \tag{3.1}$$

which can be written as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \mathsf{R} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \tag{3.2}$$

with the inverse transformation

$$x = x\cos\phi - y\sin\phi \qquad\qquad y = x\sin\phi + y\cos\phi \tag{3.3}$$

which can be written as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \mathsf{R}' \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \tag{3.4}$$

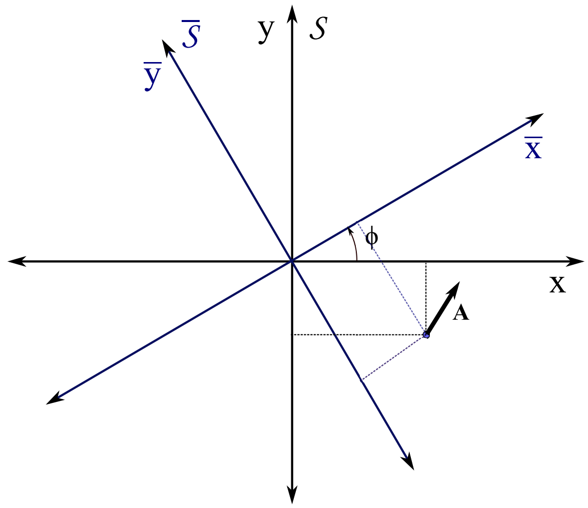

Figure 3.1: Rotating a spatial coordinate system.

the only difference being the placement of the minus sign on the sin term; this can be understood as the replacement of ϕ by $-\phi$, as the $\mathcal{S}$ system is rotated with respect to the $\bar{\mathcal{S}}$ system by angle $-\phi$. Cosine does not change sign from the change, but Sine does.)

Now we think about something more subtle: Suppose we have a vector field defined in system $\mathcal{S}$. For example, in Fig. 3.2 we show a vector $\mathbf{A}$ defined at the point P. Vector $\mathbf{A}$ has components (A_x, A_y) as measured in system $\mathcal{S}$.

If P has coordinates (x, y), we know how the people in system $\bar{\mathcal{S}}$ will get the coordinates of P, but what will they say about the components of $\mathbf{A}$? Again, it is easy to show that in system $\bar{\mathcal{S}}$ the components of $\mathbf{A}$ are

$$\bar{A}_x = \cos\phi A_x + \sin\phi A_y \qquad\qquad \bar{A}_y = -\sin\phi A_x + \cos\phi A_y$$

which is that same transformation that goes with the *coordinates*. It's easy to overlook the importance of this fact; vector $\mathbf{A}$ *is* given by two numbers in $\mathcal{S}$ but it is more than that. The term *vector* if the physics/math means more than just a set of numbers; it carries with it a method for getting the components in a rotated coordinate system. The statement that "$\mathbf{A}$ has components $(2, 3)$" does not tell you how to find the components in system $\bar{\mathcal{S}}$; the atatement "$\mathbf{A}$ has components $(2, 3)$ and $\mathbf{A}$ is a vector" *does*.

That's what "vector" really means. Of course in any given system it *is* still a set of numbers, but beyond the very first year of physics one needs to appreciate the full meaning of this word.

3.3 Invariants

In two dimensions, we have the inner product (or "dot" product),

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y = AB\cos\theta$$

θ being the angle between vectors $\mathbf{A}$ and $\mathbf{B}$, and the magnitude of a vector,

$$A = ||\mathbf{A}|| = \sqrt{A_x^2 + A_y^2} = \sqrt{\mathbf{A}\cdot\mathbf{A}}$$

But we can now appreciate a feature of the dot product and vector magnitude: They are the same whether we evaluate them using the components in the $\mathcal{S}$ system or the $\bar{\mathcal{S}}$ system. They are thus said to be **invariant** quantities. This property follows from the fact that the matrix in Eq. 3.4 is **orthogonal**, that is it has the property that $\mathsf{R}^T\mathsf{R} = \mathbf{1}$. Then, using the notation of linear algebra, for any two (true) vectors $\mathbf{x}$ and $\mathbf{y}$,

$$\bar{\mathbf{x}}\cdot\bar{\mathbf{y}} = \bar{\mathbf{x}}^T\bar{\mathbf{y}} = (\mathsf{R}\mathbf{x})^T(\mathsf{R}\mathbf{y}) = \mathbf{x}^T\mathsf{R}^T\mathsf{R}\mathbf{y} = \mathbf{x}^T\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$$

3.4 Four–Vectors

Now we use the *language* of rotations from our two-dimensional geometry to put the Lorentz transformation of Eq. 4.1 into a similar form. The Lorentz transformation is a *sort* of rotation, where instead of two space coordinates mixing, a space coordinate mixes up with the time coordinate.

We begin by putting the space and time coordinates together in one mathematical object; the space coordinates (x, y, z) will be called (x^1, x^2, x^3), where the raised indicies are written that way for a reason, and should not be confused with powers. (When we do have to square something later on, we'll be very clear about it; rarely do we have to cube something.) We make the time coordinate consistent with this scheme and having the same units by defining.

$$x^0 \equiv ct \tag{3.5}$$

and then group all four numbers as a single entity to be called simply x:

$$x \equiv (x^0, x^1, x^2, x^3)\ , \tag{3.6}$$

to be called the space–time four–vector. Usually we won't confuse the *four–vector* x and the original x coordinate, but clearly in this game we need to be sure what the symbols mean!

With this notation we rewrite the Lorentz transformation as

$$\bar{x}^0 = \gamma x^0 - \frac{\gamma v}{c}x^1 \qquad \bar{x}^1 = -\frac{\gamma v}{c}x^0 + \gamma x^1 \qquad \bar{y} = y \qquad \bar{z} = z \tag{3.7}$$

The first thing to recognize is that Eqs. 3.7 can be written more compactly as a matrix equation:

$$\begin{pmatrix} \bar{x}^0 \\ \bar{x}^2 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{3.8}$$

which can be shortened to

$$\bar{x} = \Lambda x$$

as long as we understand what the symbols mean! Here, $\bar{x}$ and x are four–vectors and Λ is a *matrix*, given in Eq. 3.8.

When the relative velocity of the two frames (the "boost") is in the x direction, as we have chosen it, the y and z coordinates don't mix with the others and all the interesting business goes on in the upper left corner of the matrix Λ.

The inverse transformation is

$$\begin{pmatrix} x^0 \\ x^2 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\frac{v}{c} & 0 & 0 \\ \gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{x}^0 \\ \bar{x}^1 \\ \bar{x}^2 \\ \bar{x}^3 \end{pmatrix} \tag{3.9}$$

or

$$x = \Lambda' \bar{x}$$

Multiplying the matrices Λ' and Λ we find

$$\begin{aligned} \Lambda'\Lambda &= \begin{pmatrix} \gamma & \gamma\frac{v}{c} & 0 & 0 \\ \gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 & 0 \\ 0 & \gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

This is a result which was to be expected, since Λ takes us from $\mathcal{S}$ to the $\bar{\mathcal{S}}$ system and Λ' takes us back, the overall effect being no change, i.e. the action of the unit matrix. So we've shown

$$\Lambda' = (\Lambda)^{-1}$$

Now, Λ is not an orthogonal matrix (since its transpose does not give its inverse) so it will *not* preserve the norm of a 4–dimensional vector– at least not in the sense of the "norm" as used above. But it does have the peculiar property that a change in sign of two of its elements does give the inverse and because of this it turns that it will preserve *something like* the norm used in regular geometry.

3.4.1 Relativistic Invariant

For a space–time point (x^0, x^1, x^2, x^3), Consider the sum/product of terms given by

$$(x^0, -x^1, -x^2, -x^3) \cdot \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$$

We can organize our work a little better using a matrix we'll call g, with

$$\begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = g \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

and then see how $x^T gx$ transforms. Evaluating in the $\bar{\mathcal{S}}$ frame, we have:

$$\bar{x}^T g\bar{x} = (\Lambda x)^T g(\Lambda x) = x^T \Lambda g \Lambda x = x^T(\Lambda g \Lambda)x \tag{3.10}$$

where we've used $\Lambda^T = \Lambda$. But we get something interesting when we calculate $\Lambda g \Lambda$:

$$\begin{aligned}
\Lambda g \Lambda &= \begin{pmatrix} \gamma & \gamma\frac{v}{c} & 0 & 0 \\ \gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \gamma & \gamma\frac{v}{c} & 0 & 0 \\ \gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ +\gamma\frac{v}{c} & -\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 & 0 \\ 0 & -\gamma^2\left(1-\frac{v^2}{c^2}\right) & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = g
\end{aligned}$$

Putting this into 3.10 gives

$$\bar{x}^T g\bar{x} = x^T gx \tag{3.11}$$

So we have shown (for boost allong the x axis, at least) that for a lorentz–type change of coordinates we have *some* property of a four-vector that stays the same between systems but it is not the everyday norm $x^T x$. It is only a tad more complicated, $x^T gx$; to find it, don't simply add the squares of the components together, add them with minuses in front of the space components.

Now if the quantity $x^T gx$ then surely $-x^T gx$ is as well, and in our notation with g, this would come from

$$\begin{aligned}
-x^T gx &= -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 \\
&= (x^0 x^1 x^2 x^3) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}
\end{aligned}$$

so that some guys define:

$$g_{\text{some guys}} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}$$

And unfortunately one sees both of these choices in the literature of physics. Usually in each book or article the author will announce his choice and then, to follow his math you just live with it; the choice of which g to use will determine many of the signs in the equations but in the end it is arbitrary; working with relativistic invariants is important for getting physics results but those results don't depend on the choice.

The choice of signs for g is known as the choice of "metric".

Some notation: As we will be finding the invariant of a vector so often, a short notation for it is essential. At the risk of confusion, we will write

$$x^2 \equiv (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{3.12}$$

3.5 Four–Vectors

We are now prepared to give a proper definition of a *relativistic* vector.

We recall that the definition of a *rotational* vector (in 2D at least) was an object which followed the same transformation law as the coordinates. Similarly,

•. If a is a four–vector, then if in frame $\mathcal{S}$ it has the components (a^0, a^1, a^2, a^3) then in frame $\bar{\mathcal{S}}$ it has components

$$\bar{a}^0 = \gamma(a^0 - \frac{v}{c}a^1) \qquad \bar{x}^1 = \gamma(a^1 - \frac{v}{c}a^0) \qquad \bar{a}^2 = a^2 \qquad \bar{a}^3 = a^3 \tag{3.13}$$

or $\bar{a} = \Lambda a$.

Furthermore, for any four–vector a, the quantity

$$a^2 = (a^0)^2 - (a^1)^2 - (a^2)^2 - (a^3)^2$$

will be the same if evaluated in any reference frame.

From this, one can show that if a and b are four–vectors, then the combination

$$a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3$$

is also invariant. This combination is analagous to the dot product in regular vector algebra, and the notation we use is similar:

$$a \cdot b = a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3$$

Now for some more notation. *With the choice of metric we have*, we will define the components of a vector associated with x,

$$a_0 \equiv a^0 \qquad a_1 \equiv -a^1 \qquad a_1 \equiv -a^1 \qquad a_2 \equiv -a^2 \qquad a_3 \equiv -a^3 \tag{3.14}$$

Technically, people say that x^μ, with $\mu = 0, \ldots 3$ is the **covariant** form of the vector x and x_μ is the **contravariant** form. (Note, for the other choice of the g matrix, we would define a contravarient x_μ with only a minus sign on the x^0 coordinate.)

With the definition in 3.14 the invariant is

$$(a^0)^2 - (a^1)^2 - (a^2)^2 - (a^3)^2 = \sum_{\mu=0}^{3} a^\mu a_\mu$$

If fact, we will often shorten this to a^2; it's a little risky to write that down, as one might confuse it with the y spatial component of the four-vector a. Generally we can tell from the context what is meant, but it is always true that when we simplify the notation in physics we run the risk of sometimes forgetting the meaning!

And now *another* bit of notation. We will be performing a sum over coordinates so often that when we see a product of vector components with the same indices where one index is up and the other is down, we will agree that the index is to be summed over (from 0 to 3). Thus

$$a^\mu b_\mu \qquad \text{means} \qquad \sum_{\mu=0}^{3} a^\mu b_\mu$$

If it ever happens that we *don't* want to do this sum, we can explicitly say so, but the savings in time from not writing down the summation symbol each time is worth it.

We can write the Lorentz transformation in the new notation:

$$\Lambda^{\text{row}}_{\text{column}} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Then

$$\bar{x}^\mu = \Lambda^\mu_{\ \nu} x^\nu$$

The awful transformation equations back in Eq. 4.1 now have a very simple form, but there are lot of details hiding behind this simple notation!

While we're on the subject of compact notation, we recall that the relation

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = (g) \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$$

can be written as

$$x_\mu = g_{\mu\nu} x^\nu \tag{3.15}$$

with our summation convention understood. The matrix elements of g can be written with the indicies up or down, with:

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

So the indices of any vector can be lowered by applying the g matrix.

3.6 Relativistic Tensors

There are also quantities with two indices (whose values go from 0 to 3). The quantity $T^{\mu\nu}$ contains sixteen different values. If it is a **tensor** then we know what those values are in a different reference frame; they come from the following formula, which is much like the one for vectors but with the Λ matrix applied twice. In frame $\bar{S}$ the values are

$$\bar{T}^{\mu\ \nu} = \Lambda^{\mu}_{\ \rho}\Lambda^{\nu}_{\ \sigma}T^{\rho\sigma} \tag{3.16}$$

Such quantities won't be covered for a while, but we will come back to them when we study electromagnetism and relativity because the E and B fields are not elements of relativistic *vectors*, rather they are both components of a single relativistic *tensor*.

Chapter 4

Examples and Applications of Lorentz Transforms

Before we go too far with the mathematics of the Lorentz transforms and the invariants it will be useful to consider *a few* implications of the Lorentz transformation and the invariants

for reference again, the Lorentz transformation equations for a frame $\bar{\mathcal{S}}$ moving in the $+x$ direction with speed v with respect to frame $\mathcal{S}$ are:

$$\bar{x} = \gamma(x - vt) \qquad \bar{y} = y \qquad \bar{z} = z \qquad \bar{t} = \gamma\left(t - \frac{v}{c^2}x\right) \tag{4.1}$$

with the reverse transformation coming from solving these for (x, y, z, t) or more simply changing the sign of v:

$$x = \gamma(\bar{x} + v\bar{t}) \qquad y = \bar{y} \qquad z = \bar{z} \qquad t = \gamma\left(\bar{t} + \frac{v}{c^2}\bar{x}\right) \tag{4.2}$$

4.1 Time dilation

We take the example of a rapidly–moving particle which has a lifetime τ when measure *in its rest frame*.

Armed with the mathematics of *relativistic* vectors we can now think of developing the laws of motion in way so that they have the same form in all reference frames. That's the goal: The fundamental laws of physics have the same mathematical form in all frames!

Chapter 5

Relativistic Dynamics

For our purposes, Newton's 2nd law for the motion of a *particle* of mass m is now best expressed as

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \qquad \text{where} \qquad \mathbf{p} = m\mathbf{u} \ .$$

We have to come up with something new, but at small velocities it has to reduce down to this form.

In this section I won't *prove* the relativistic version of Newton's 2nd law; it's not possible to show how a more general law follows from the limited one that we already have. But we can make it plausible from the mathematical aspects of the Lorentz transformation seen in the last chapter. In the end the validity of the law is established by experiment.

5.1 Relativistic Momentum

We might expect to have a law which contains relativistic *vectors* so that its transformation properties will be simple. If we take the mass m to be a property of a particle regardless of its state of motion (and eventually we will use the term "mass" that way) then as we have seen, the velocity of a particle and also the product $m\mathbf{u}$ can't represent the spatial part of a relativistic vector, because as we saw, even the y and z components of $\mathbf{u}$ change under a Lorentz transformation of the coordinates.

Equivalently, we can consider the definition of velocity, $\frac{d\mathbf{x}}{dt}$, thinking of the derivative as the limit of a ratio of finite differences. While the numerator $d\mathbf{x}$ does indeed transform like a vector it is divided by a time difference which also transforms. The result is not (part of) a Lorentz vector.

But there is a way out! As the particle moves over a spatial interval $d\mathbf{x}$, the corresponding time interval dt depends on the reference frame, but the time interval *in the frame of the moving particle* is something everyone can at least *agree* on. If in our frame we measure a time interval dt, in the frame of the particle (where $d\bar{x} = 0$) the time interval will be

$$d\tau = \gamma dt = \frac{dt}{\sqrt{1 - \frac{u^2}{c^2}}} \tag{5.1}$$

which is larger than dt. A time interval measured in the reference frame of a moving particle is called the **proper time** interval. While we don't want to call the new ratio $\mathbf{v}$, we can called it η and define

$$\eta \equiv \frac{d\mathbf{x}}{d\tau} = \frac{1}{\sqrt{1-u^2/c^2}}\frac{d\mathbf{x}}{dt} = \gamma\mathbf{u}$$

where $\gamma = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$. We note that η is a peculiar kind of quantity, a sort of hybrid in that $d\mathbf{x}$ is measured in our frame while $d\tau$ comes from a measurement in the frame of the particle.

Notice that while previously we have $\mathbf{v}$ stand for the velocity of a reference frame and $\mathbf{u}$ the velocity of a particle, here they are the same thing so we will simply use γ to stand for $\frac{1}{\sqrt{1-u^2/c^2}}$.

With this definition of η, we can be assured that a relativistic three–momentum defined

$$\mathbf{p} \equiv \frac{m\mathbf{u}}{\sqrt{1-u^2/c^2}} = \gamma m\mathbf{u} \tag{5.2}$$

could be part of a relativistic vector, if we can only find a suitable 0^{th} component!

The zeroth component is (essentially) the relativistic *energy*.

5.2 What People Used to Say, But We Won't

When we look at expression 5.2 we see our familiar definition of momentum with an extra factor of $\sqrt{1-u^2/c^2}$ downstairs which is close to 1 for small velocities but which will make a significant difference for speeds comparable to c.

What is the best way to understand this new factor?

Currently the thinking is more or less the way we arrived at it, that is, it is a factor which is needed to fix the *velocity vector* (through the time factor) so that we get a vector. That's not simple but that's what we did!

Not so long ago it was common to say that this factor of $\sqrt{1-u^2/c^2}$ was an adjustment to the *mass* and so one could say that in relativity the mass of a particle effectively changed with velocity. What people would often say is that a particle has a *rest mass*, to be written as m_0 while a moving particle has a mass

$$m(u) = \frac{m_0}{\sqrt{1-u^2/c^2}} = \gamma m \qquad \text{(Old useage!)}$$

While there's nothing really *wrong* with such a statement it is now thought to be unnecessary and there's little to be gained from this idea and possibly not worth the confusion over the meanings of the term "mass" which might take place. As we'll see, this velocity-dependent mass is basically the relativistic *energy*. So for us and in most books nowadays there is *one* "mass" and it is the thing that was once called the "rest mass" m_0.

5.3 Relativistic Energy

The expression for the relativistic energy of a freely–moving particle of mass m (no potentials) is given by

$$E = \frac{mu^2}{\sqrt{1 - u^2/c^2}} = \gamma m u^2$$

We will justify our expression for the energy by considering (what ought to be) the *increase* in energy if we do work on a particle; recall that is how we arrive at the definition of the Newtonian kinetic energy $K = \frac{1}{2}mv^2$:

$$W_{\text{net}} = \int_1^2 \mathbf{F}_{\text{net}} \cdot d\mathbf{r} = \int_1^2 m\frac{d\mathbf{v}}{dt} \cdot d\mathbf{r} = \int_1^2 m\frac{d\mathbf{v}}{dt} \cdot \frac{d\mathbf{r}}{dt}\, dt$$

Use

$$\frac{d\mathbf{v}}{dt} \cdot \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{v}}{dt} \cdot \mathbf{v} = \tfrac{1}{2}\frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) = \tfrac{1}{2}\frac{d(v^2)}{dt}$$

and then we get

$$W_{\text{net}} = \int_1^2 \frac{m}{2}\frac{d(v^2)}{dt}\, dt = \tfrac{1}{2}mv_2^2 - \tfrac{1}{2}mv_1^2 = \Delta K$$

What happens in relativity? Using the same *starting* expression for the work (but a different substitution for the force $\mathbf{F}$) we now get

$$W_{\text{net}} = \int_1^2 \mathbf{F}_{\text{net}} \cdot d\mathbf{r} = \int_1^2 \frac{d\mathbf{p}}{dt} \cdot d\mathbf{r} = \int_1^2 \frac{d\mathbf{p}}{dt} \cdot \frac{d\mathbf{r}}{dt}\, dt \tag{5.3}$$

We can work on the insides of this integral; it's a little messier than the Newtonian case! Use $\mathbf{u} = d\mathbf{r}/dt$ and the (relativistic) definition of $\mathbf{p}$ to get:

$$\frac{d\mathbf{p}}{dt} \cdot \frac{d\mathbf{r}}{dt} = \mathbf{u} \cdot \frac{d}{dt}\left(\frac{m\mathbf{u}}{\sqrt{1 - u^2/c^2}}\right) \tag{5.4}$$

The next step is a little messy since $\mathbf{u}$ shows up in two places; use the product and chain rules and note that

$$\frac{d(u^2)}{dt} = \frac{d}{dt}(\mathbf{u} \cdot \mathbf{u}) = 2\mathbf{u} \cdot \frac{d\mathbf{u}}{dt} \tag{5.5}$$

This gives

$$\begin{aligned}
\frac{d\mathbf{p}}{dt} \cdot \frac{d\mathbf{r}}{dt} &= m\mathbf{u} \cdot \left[\frac{1}{\sqrt{1 - u^2/c^2}}\frac{d\mathbf{u}}{dt} + \frac{-1}{2}(-1/c^2)\frac{\mathbf{u}}{(1 - u^2/c^2)^{3/2}} 2\mathbf{u} \cdot \frac{d\mathbf{u}}{dt}\right] \\
&= m\left[\frac{1}{\sqrt{1 - u^2/c^2}}\mathbf{u} \cdot \frac{d\mathbf{u}}{dt} + \frac{u^2/c^2}{(1 - u^2/c^2)^{3/2}}\mathbf{u} \cdot \frac{d\mathbf{u}}{dt}\right] \\
&= m\left[\frac{(1 - u^2/c^2) + u^2/c^2}{(1 - u^2/c^2)^{3/2}}\mathbf{u} \cdot \frac{d\mathbf{u}}{dt}\right] = \frac{m}{2(1 - u^2/c^2)^{3/2}}\frac{d(u^2)}{dt}
\end{aligned}$$

where we've again used 5.5 in the last step. We've now got something written entirely in terms of u^2 and it can be integrated to give a total time derivative. We get:

$$\frac{d\mathbf{p}}{dt} \cdot \frac{d\mathbf{r}}{dt} = \frac{d}{dt}\left[\frac{mc^2}{\sqrt{1-u^2/c^2}}\right]$$

This is exactly what we want; putting this back into 5.3 we get

$$\Delta E = W_{\rm net} = \int_1^2 \frac{d}{dt}\left[\frac{mc^2}{\sqrt{1-u^2/c^2}}\right] dt = \left[\frac{mc^2}{\sqrt{1-u^2/c^2}}\right]\Bigg|_1^2$$

...a very interesting result. The change in energy is the change in the quantity in the braces, which would imply that the energy itself is that quantity *plus some constant*:

$$E = \frac{mc^2}{\sqrt{1-u^2/c^2}} + E_0 = \gamma mc^2 + E_0 \tag{5.6}$$

What is the constant E_0? We note that for $u = 0$ (particle at rest) the formula gives $E = mc^2 + E_0$ and maybe we want $E_0 = -mc^2$ to give zero energy for a particle at rest. It turns out that we *don't* want this. Experiments in which particles can change identity will satisfy a conservation of energy law with $E_0 = 0$ so that the right formula is

$$E = \frac{mc^2}{\sqrt{1-u^2/c^2}} = \gamma mc^2 \tag{5.7}$$

which means that even when the particle is at rest we associate an energy, $E_{\rm rest} = mc^2$, a very famous equation. But we note that *as we are using the symbol m for the (rest) mass*, it is only true for a particle at rest. Had we gone with the notation mentioned above, the mass-dependent m, i.e. $m(v) = \gamma m$ then it would always be true. But then the two quantities $m(v)$ and E just differ by a a factor of c^2 and the two quantities are *basically* the same. (Factors of c will be seen to just get in the way of our understanding; professionals basically ignore them.)

Before considering the meaning of a rest energy, we want to see how this result relates to the energy that we knew from Newtonian mechanics; from the messy derivation above this is not clear at all!

Our previous results can only apply to the case $u \ll c$ so we write the energy as

$$E = mc^2\left(1 - \frac{u^2}{c^2}\right)^{-1/2}$$

Here we recall the binomial expansion; when $x \ll 1$ the Taylor expansion

$$(1 \pm x)^n = 1 \pm nx + \frac{n(n-1)}{2!}x^2 \pm \frac{n(n-1)(n-2)}{3!}x^3 + \ldots$$

converges rapidly and gives a good approximation when we truncate it; so here with $x = v^2/c^2$ and $n = -\frac{1}{2}$ it gives us

$$\begin{aligned} E &\approx mc^2 + \tfrac{1}{2}mc^2\frac{u^2}{c^2} + \frac{3}{8}mc^2\frac{mu^4}{c^4} + \cdots \\ &= mc^2 + \tfrac{1}{2}mu^2 + \frac{3}{8}\frac{mu^4}{c^2}\cdots \end{aligned} \tag{5.8}$$

Now we look at the individual terms in 5.8. The first term is the rest energy, something totally absent from Newtonian mechanics. The second term is the Newtonian kinetic energy, $\frac{1}{2}mu^2$. The third term (and the higher–order terms) can be taken as a correction to the Newtonian value of the kinetic energy. With that in mind, we define the *relativistic* kinetic energy T of a particle

$$T = E - mc^2 \tag{5.9}$$

which is *not* equal to $\frac{1}{2}mu^2$ but is close to that value for speeds small compared to c.

5.4 Relation Between Momentum and Energy

We now have the relativistic definitions

$$\mathbf{p} = \gamma m \mathbf{u} \qquad \text{and} \qquad E = \gamma m c^2$$

and this gives

$$\mathbf{p}^2 c^2 = \frac{m u^2 c^2}{(1 - u^2/c^2)} \qquad E^2 = \frac{m^2 c^4}{(1 - u^2/c^2)}$$

Subtracting the first from the second gives

$$E^2 - \mathbf{p}^2 c^2 = m^2 c^2 \left(\frac{c^2}{(1 - u^2/c^2)} - \frac{u^2}{(1 - u^2/c^2)} \right) = m^2 c^4 \left(\frac{1 - u^2/c^2}{1 - u^2/c^2} \right) = m^2 c^4$$

which has the same value in all frames as the mass m is a characteristic of the particle. This suggests that we can construct a relativistic four–vector out of the energy and momentum:

$$p = (p^0, p^1, p^2, p^3) = (E/c, p^1, p^2, p^3) \equiv (E/c, \mathbf{p}) \tag{5.10}$$

Chapter 6

Rants; Relativity Misinformation and Myths

I began these notes after experiencing a lot of frustration with the way relativity was presented to undergraduates in the so-called "Modern Physics" courses. In short, I got tired of the emphasis on the poles and barns, the silly examples which try to show some kind of cute paradox (actually a *non–intuitive result*) when the Lorentz transformation is used.

Now, in some ways the weight given to these silly examples is unavoidable. Students take the Modern Physics course *before* learning very much about electromagnetism (Maxwell's equations) and before hearing much about nuclear and particle experiments, so when it's time for examples of the unique predictions of special relativity the only things available are some whimsical illustrations of the new ideas of space and time. But it's also possible that if this is the major part of a student's first encounter with the subject, then it does as much harm as good. The student can come away from such examples thinking that these funny effects aren't even *real*, that they only come about as an artifact of the peculiar way they are derived, with light-clocks and sending signals back and forth through space with light waves. Or that if we only refrain from (rapidly) running around with poles we won't need to worry about relativity. As the example of the muon lifetime shows, it's *not* a trick with light-clocks and mirrors. Relativity tells *what you are going to measure*.

Relativity is not goofy. It's important. But specifically, I've always objected to the following myths:

6.1 Changing Mass in Relativity Justifies the Idea in Newtonian Mechanics

As mentioned, the idea of a speed-dependent mass of a particle is an *interpretation* of the equations and as is now thought, not a very profitable one— even if the great Feynman advocated this point of view in the famous Lectures. It's not *wrong* to talk this way, because we've always got the equations as the last word on the subject.

But some people carry this notion back to Newtonian mechanics where they say that $\mathbf{F} = \frac{d\mathbf{p}}{dt}$ is a more general way of writing Newton's law as it allows for the possibility of changing mass... and behold, in relativity we do have that changing mass! It all fits!

There is lots of confusion here.

Newton's mechanics is a theory of idealized particles. Though particles can interact (including processes where they can stick and un-stick) their masses *do not change*; mass is conserved in Newtonian mechanics. Thus $F = dp/dt$ and $F = ma$ have the same content. While *the form* of $F = dp/dt$ goes over to relativity while $F = ma$ does not, it is *arbitrary* to say this is due to a changing mass; one could just as well say it is due to a replacement of the velocity factor, or a better expression for p as a whole.

$F = dp/dt$ in Newtonian mechanics should be taken as *one way* of expressing Newton's 2^{nd} law, one which can be taken over to the improved theory. It sometimes happens in physics that one of the *equivalent* formulations of a theory allow us to go to the "next level" while the others must stay behind.

6.2 $E = mc^2$ Clued Us In to Nuclear Energy

It is sometimes said (or implied) that this formula led to the development of nuclear energy.

The energy value of an isolated bit of matter at rest, $E = mc^2$ is applicable to *all* matter, whether or not we understand its internal "gears and wheels" and whether those gears and wheels come from strong or electromagnetic forces. It applies to chemical energy as well as nuclear energy.

It is true that nuclear binding energies are large enough (at the level of a few percent of the nuclear rest mass) that they can be measured fairly readily, while chemical binding energies are very hard to determine from mass measurements. But the masses of individual atoms don't *exactly* add up to give the mass of the molecule they form.

Furthermore, since *mass is conserved* in Newtonian mechanics you can simply add mc^2 to the energy of a particle of mass m and you will still have a valid law for the conservation of energy. (In Newton's mechanics only the *differences* in energy have any meaning.)

It's just that in Newtonian mechanics there's *no reason* to go adding mc^2 to the energies.

And this is all related to...

6.3 Because of $E = mc^2$, Normal Matter at Rest "Contains" Vast Amounts of Energy

The claim is often made that if some tiny speck of matter could be "completely converted to energy, there would be enough energy to..."

There is really no distinction between "matter" and "energy". There are only *particles*. There are of course particles which do have a non-zero rest mass and ones for which the mass is too small to be detected. The latter is probably what people mean when they say "energy", but under the right conditions *any* particle can be given enough energy so that its rest mass contribution is negligible, electrons in Deep Inelastic Scattering being an example which comes to mind.

Now it is true that some particles with a rest mass *can* change into particles with none, as in the case of the decay of the π^0 to two photons. But that isn't the case with very small bits of normal matter which by usual (!) nuclear processes can never lose a large part of their rest mass. Basically you can't talk about regular matter being entirely converted to "energy" because *it can't happen*.

6.4 Various Rewritings of History

The biggest misconception here is that Einstein immediately floored all physicists with his ideas about transformations of space and time so that the old ideas were tossed out immediately. Reading the contemporary textbooks by some of the secondary geniuses of the time shows this was not the case.

In 1908 Lorentz wrote a set of notes on the status of the "theory of electrons", meaning the theory of charged particles and the electromagnetic field as it was understood at the time. This was long enough after Einstein's papers of 1905 that his work was well known, but his basic strategy was not accepted.

In 1908 Lorentz mentions Einstein *postulating* the space-time transformation but with (it seems) a little bafflement and frustration. Of course, Lorentz had the space–time transformation rule in hand before Einstein (that's why they're named after him) but he felt that they had to be *derived* from the processes which *made* the electromagnetic field or else one didn't truly understand them. Simply *asserting* them seemed like a cop-out. And in fact Lorentz did a fairly convincing job of deriving them from the supposed motion of macroscopic objects through an aether. But Lorentz always treated the new coordinates in the new frame as *effective* coordinates, not the real thing. And that was how things were in 1908.

In 1915 he revised this book and by that time it was clearer that Einstein had found a very deep principle of reality. In the new edition of the notes issued that year, Lorentz is verbally kicking himself for treating the $\bar{t}$ coordinate as (only) *something like* a real time coordinate and admits that he would have given much more space to Einstein's ideas if he were re-writing the whole book.

In 1920 Arthur Eddington (the British astronomer famous for his abuse of Chandresekhar) wrote a basic summary of Einstein's theories of special and general relativity. In it he raves about the contribution of Einstein in ridding us of unnecessary assumptions in physics, such as the way to transform coordinates between reference frames.

Surely the great Sir Arthur himself had been cured of all *his* false assumptions, right?

Sadly, no. After remarking on how one could no longer assign any obvious properties to the aether, and how it plays no useful role in any theory, Eddington says that we can't simply *give up* the aether!

> "Some would cut the knot by denying the aether altogether. We do not consider that desirable, or, so far as we can see, possible; but we do deny that the aether need have such properties as to separate space and time in the way supposed."

So to some degree the aether was alive and well in 1920[1].

[1] To me, this seems a little odd given the simultaneous advances in atomic theory which gave a picture of light as discrete bundles of energy (photons) moving though space. This wouldn't seem to require an "aether" picture.