Relativity

James Emery

Last Revision: 4/7/2017

Contents

1 Introduction

There are many books on special relativity. I particularly like a book by Robert Katz titled, **An Introduction to the Special Theory of Relativity**.

Among the many books on the General Theory of Relativity the books I especially like are **A First Course in General Relativity** by Bernard Schutz, **Gravitation and Spacetime** by Ohanian and Ruffini, **Introduction to the theory of Relativity** by Peter Begmann, and **General Relativity**, by Robert Wald. There are many other good books listed in the bibliography.

Einstein's first paper on relativity has the title, **On the Electrodynamics of Moving Bodies.** His original motivation was not in mechanics, but rather in the curious fact that the existence of Electric and Magnetic fields generated by magnets and conductors, seems to depend upon which object is at rest. Not just on relative velocity, and so seems to contradict the fundamental idea of Newtonian Mechanics that there is no absolute coordinate system that is at rest. Special Relativity Theory is now usually first presented in mechanics, with little mention of electrodynamics. Much of the mathematics of special relativity is simple. This is not so for The General Theory of relativity. It is a theory of Gravitation that explains gravity as a distortion of space-time itself in the neighborhood of mass. This theory involves difficult ideas from Differential Geometry and Tensor Analysis. Compared to this theory, the mathematics of the introductory part of the special theory is rather simple. Although, because we lack intuition for this special theory, simple problems can be hard to solve and interpret. Einstein received mathematical assistance from his friend Michele Angelo Besso (born May 25, 1873 Riesbach - died March 15, 1955 Genova). Assistance was also given by Marcel Grossmann. Marcel Grossmann (born in Budapest on April 9, 1878 - died in Zurich on September 7, 1936) w as a mathematician, a friend, and a classmate of Albert Einstein. He became a Professor of Mathematics at the Federal Polytechnic Institute in Zurich, today the ETH Zurich, specializing in geometry. It was Grossmann who emphasized the importance of a non-Euclidean geometry to Einstein, which was a necessary step in the development of Einstein's general theory of relativity. The Abraham Pais biography of Einstein, **Subtle is the Lord**, suggests that Grossman mentored Einstein in the necessary tensor theory. This book contains much scientific material.

2 The Special Theory

The fundamental assumptions that lead to relativity theory, are that (1), the velocity of light is to be constant in all inertial coordinate systems, and (2), the equations of physics are covariant. This means that they take the same form in all inertial coordinate system. It turns out that space-time is four dimensional. It consists of three spacial coordinates, and one time coordinate. The intuitive idea, that time is the same everywhere, is false. A pair of simultaneous events in one inertial system may not be simultaneous in another inertial system. An event in space time, is like a point in classical space, but is 4-dimensional, including not just the three space coordinates, but also the time. In a given fixed inertial system we measure as usual. We measure length with meter sticks, and time with physical clocks. But if light is to have constant velocity in a pair of coordinate frames that are moving relative to each other, then the time coordinate can not be taken as the same in both coordinate systems. The velocity of light c is taken to be constant, so observers may make measurements of distant events by using this constant light velocity. For example a radar device can be used to measure the coordinates of a pair of events.

Let v be the relative velocity between two coordinate frames. For simplicity assume this velocity is in the x direction. The two frames are labeled as an unprimed frame and a primed frame. Let c be the velocity of light,which is approximately $c = 3 \times 10^8$ m/s. We define

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

Then

$$\gamma \geq 1.$$

As v gets close to c, γ gets large. For example, if $v = .995c$ then γ is about 10, if $v = .9c$ then γ is between 2 and 3. At normal velocities, γ is very close to 1. Suppose the laboratory coordinate system is denoted with unprimed coordinates, and that the system moving with velocity v relative to the laboratory is denoted with primed coordinates.

The Lorentz Transformation relating the two coordinate systems is

$$x' = \gamma(x - vt)$$

$$y' = y$$

$$z' = z$$

$$t' = \gamma(t - \frac{xv}{c^2}),$$

where v is the velocity of the primed system with respect to the unprimed system. The inverse Lorentz transformation is

$$x = \gamma(x' + vt')$$

$$y = y'$$

$$z = z'$$

$$t = \gamma(t' + \frac{x'v}{c^2}),$$

where we reverse the sign of v.

The Lorentz Transformation is derived by using the constancy of the velocity of light in the two systems. So if a light wave moves out from a point source, it moves out in circles in both coordinate systems. This fact can be used to derive the Lorentz transformation, (see for example **Eisberg**). To derive the transformation one assumes that the transformation is linear, and that photons move in straight lines in each coordinate system. Alternately the transformation is derived in the **Krane** book using a special clock, consisting of a device of length L that emits particles at one end, and reflects them from a mirror on the other end. In a following section we go through the derivation of the Lorentz transformation following the derivation given in Peter Bergmann, **Basic Theories of Physics.**

The interval

$$\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2,$$

is preserved by the Lorentz transformation. This is the metric between events, the metric in space-time. We can verify this preservation directly by a simple substitution. So

$$c^2t'^2 = \gamma^2 c^2(t - \frac{vx}{c^2})^2$$

$$= \gamma^2(ct - \frac{vx}{c})^2$$

$$= \gamma^2(c^2t^2 - 2vxt + \frac{v^2x^2}{c^2})$$

and similarly

$$x'^2 = \gamma^2(x^2 - 2xvt + v^2t^2).$$

So

$$c^2t'^2 - x'^2 =$$

$$\gamma^2(c^2t^2 - x^2 - \frac{v^2}{c^2}(c^2 - x^2))$$

$$= \gamma^2(1 - \frac{v^2}{c^2}(c^2 - x^2)$$

$$= c^2t^2 - x^2.$$

The values of the y and z coordinates do not change under the transformation, so the metric distance is preserved,

$$c^2t'^2 - x'^2 - y'^2 - z'^2 = c^2t^2 - x^2 - y^2 - z^2.$$

Let the difference of coordinates between two events be t, x, y, z. Let L be the spacial distance between events so that

$$L^2 = x^2 + y^2 + z^2.$$

If the metric

$$c^2t^2 - L^2$$

is positive, then the events are called time-like and we define a proper time T between the events by

$$T^2 = t^2 - L^2/c^2 = t'^2 - L'^2/c^2.$$

This proper time is the same in all coordinate systems. It is called proper time because we can find a coordinate system in which the spacial distance between the events is zero. Then the events occur at the same spacial point at successive times. To find such a coordinate system, suppose a primed coordinate system moves at velocity v with respect to the original coordinate system with the x axis passing through the two events. Then choose the velocity of the primed coordinate system so that

$$v = \frac{x}{t}.$$

Then the Lorentz transformation gives us

$$x' = \gamma(x - vt) = 0,$$

so that $L' = 0$. And so

$$T^2 = t'^2 - L'^2/c^2 = t'^2.$$

So that in that primed coordinate system, the proper time corresponds to the actual time. In any other coordinate system, for time-like events, the proper time is larger than the time between these events. So in the twins paradox where a twin leaves earth in a rocket ship and travels at a speed at near the velocity of light when he returns he is younger than his twin who remained on earth. This is true because the time interval experienced by the stay at home twin between the two events of his twin leaving and returning is his proper time interval, and so must be greater than the time interval measured by the traveling twin. This difference can be any amount depending on how close the rocket ship speed is to c. We speak both of a proper time interval, and a proper length interval between events, depending upon whether the events are time-like or space like. So for a time-like interval, a proper time interval for two events is the time measured at a point fixed in a given coordinate system. Similarly, if the metric

$$c^2t^2 - L^2$$

is negative, the events are called space-like, and we define a proper length

$$S^2 = L^2 - c^2t^2 = L'^2 - c^2t'^2.$$

Proper length is a length measured between two points at a fixed time in a given coordinate system. If two events are time-like, then a signal sent from the first event can reach the second event. If two events are space-like, then a signal sent from the first event can not reach the second event. If we take units in which the velocity of light is 1, then the slope of a line joining the two events in a spacetime diagram will be greater than one, for a space-like interval the slope will be less than one.

By taking differences, we see that a time interval in the primed system dt' is transformed to

$$dt = \gamma(dt' + \frac{dx'v}{c^2})$$

in the unprimed system. For a clock at a fixed point in the moving system $dx' = 0$. So we have

$$dt = \gamma dt'.$$

We see that we have **Time Dilation.** A clock fixed in the primed system runs slow when measured in the unprimed system. So suppose $v = .995c$ and γ is about 10. If the "tick tock" in the primed system is 1 second, then in the unprimed laboratory system it will be about 10 seconds. Again, by taking differences, and using the Lorentz transformation, length intervals in the two coordinate systems satisfy

$$dx' = \gamma(dx - vdt).$$

If the ends of a rod, are fixed in the moving primed coordinate system, and are measured in the laboratory system simultaneously, so that $dt = 0$, then

$$dx = \frac{dx'}{\gamma}.$$

So if we measure two ends of a rod in the unprimed laboratory system, which is fixed in the moving system, and has length dx' in that system, we find that we get **Length Contraction:,** $dx < dx'$. A meter stick fixed in the moving system, will be measured as less than one meter in the laboratory system. At about $v = .995c$ the measured length will be about 1/10 meter.

Let u be the velocity of a body in the unprimed frame and u' the velocity in the primed frame. Velocity transforms as follows

$$u' = \frac{u - v}{1 - uv/c^2}.$$

This follows by dividing dx' by dt' (We shall take liberties with mathematical rigor). We define the rest mass m_0 of a body in the usual way by comparing to standard masses such as the standard Kilogram located in Paris, when the body is at rest in our coordinate system.

In the classical theory of the electromagnetic field it is shown that momentum is conserved provided that a proper momentum is assigned to the electromagnetic field itself. Thus the mass of a charged particle varies with its speed because of the electromagnetic momentum assigned to the field surrounding it. Considerations such as this lead to the idea that momentum can be conserved provided a particle of rest mass m_0 is assigned a momentum

$$p = \gamma m_0 v,$$

rather than as the classical value

$$p = m_0 v.$$

Radiation must have a momentum. The momentum of an electromagnetic field can be derived by noting that a quantity of radiation has energy density given by

$$\frac{\mathbf{B} \cdot \mathbf{H}}{2} + \frac{\mathbf{E} \cdot \mathbf{D}}{2},$$

And that electromagnetic radiation flows through a surface with intensity given by the Poynting vector

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}.$$

(See Richtmyer et al, pp 65-67 and Appendix I.) So an electromagnetic field has momentum and exerts radiation pressure. So one would have assigned a mass to the electromagnetic field, because classically in order to have momentum, one must have mass. But radiation consists of photons, which are massless. By this we mean that

$$m_0 = 0.$$

So we assign just momentum and energy to the field, not mass. Considering how velocities transform, we could not have conservation of momentum in both coordinate frames if we use the classical definition of momentum. Stated another way equations involving momentum are covariant with this definition. And of course

$$p = \gamma m_0 v,$$

works, experiments show this, and this is the ultimate justification for any theory in Physics. Sometimes this is written as

$$p = mv,$$

where

$$m = \gamma m_0.$$

One would say that as a body is accelerated to velocity v, its mass increases from its rest m_0 to m. However, I think that this variable mass point of view, is no longer popular among physicists. When we talk about mass, we usually mean the rest mass. Defining

$$p = mv,$$

where

$$m = \gamma m_0,$$

is not a very useful definition of momentum in the case of a particle with zero rest mass. It is possible to eliminate the occurance of the velocity and get a more general formula. We shall do that in a moment. Suppose a body has a rest mass m_0, and suppose we accelerate it with a force F. By Newton's second law the force is

$$F = \frac{dp}{dt} = \frac{m_0 dv/dt}{(1 - v^2/c^2)^{3/2}}.$$

The kinetic energy gained by a particle which is acted on by a force F, is

$$K = \int F \cdot dr = \int_0^t \frac{m_0 dv/dt}{(1 - v^2/c^2)^{3/2}} v dt$$

$$= \left[\frac{m_0 c^2}{(1 - v^2/c^2)^{1/2}} \right]_0^t = \gamma m_o c^2 - m_0 c^2.$$

The quantity m_0c^2 is called the rest energy of the particle. This result suggests that we take the total energy of a particle to be

$$E = \gamma m_0 c^2.$$

Again if we define

$$m = m_0\gamma,$$

then we get the famous formula

$$E = mc^2.$$

This formula was presented by Einstein in a second 1905 paper on Relativity, **Ist die Tragheit eines Korpers von sienem Energiegehalt abhangig?**, (Translation: Does the Inertia of a Body Depend Upon its Energy-Content?). In that paper he used an argument dealing with a body that emits radiation. We have

$$p^2 = \gamma^2 m_0^2 v^2 = (\gamma^2 m_0^2 c^4)\frac{v^2}{c^2} = E^2\frac{v^2}{c^4}.$$

Now here is the more general formula not involving v. We have

$$E^2 = m_0^2\gamma^2 c^4.$$

So

$$E^2(1 - v^2/c^2) = m_0^2 c^4.$$

Then

$$E^2 = m_0^2c^4 + E^2\frac{v^2}{c^2} = m_0^2c^4 + c^2E^2\frac{v^2}{c^4}$$
$$= m_0^2c^4 + c^2p^2.$$

This is a more fundamental formula than $E = mc^2$, and holds for particles with zero rest mass. One can remember this as a kind of Pythagorean formula, where the hypotenuse is the energy E, one leg of the triangle is m_0c^2 and the other is cp. Thus

$$E^2 = (m_0c^2)^2 + (cp)^2.$$

For a photon, by letting m_0 go to zero, we get the expression

$$E^2 = c^2p^2.$$

From the special theory of relativity we have the expression for energy of a particle (see for example Emery **Relativity** relativ.pdf, or a book on relativity or modern physics)

$$E^2 = (m_0c^2)^2 + (cp)^2,$$

where m_0 is the rest mass. For a photon, by letting m_0 go to zero, we get the expression

$$E^2 = c^2p^2,$$

or

$$E = cp,$$

According to Einstein's theory of the photoelectric effect we have its energy equal to

$$E = h\nu,$$

where ν is the frequency. Thus

$$h\nu = cp$$

so

$$h = \frac{c}{\nu}p = \lambda p$$

because

$$\nu\lambda = c$$

where λ is the wavelength. So the wavelength of the photon is

$$\lambda = \frac{h}{p}$$

De Broglie applied this to all particles getting a wavelength λ for every particle of momentum p,

$$\lambda = \frac{h}{p}$$

According to Einstein's theory of the photoelectric effect the energy of the photon is equal to

$$E = h\nu,$$

where ν is the frequency. Thus

$$h\nu = cp$$

so

$$h = \frac{c}{\nu} p = \lambda p$$

because

$$\nu \lambda = c$$

where λ is the wavelength. So the wavelength of the photon is

$$\lambda = \frac{h}{p}$$

DeBroglie applied this to all particles so getting a wavelength for every particle of momentum p, of

$$\lambda = \frac{h}{p}$$

Of course the idea of the massless photon comes from the theory of the photoelectric effect and quantum mechanics.

The expression for the kinetic energy becomes the classical expression for small velocities. Indeed, suppose a particle is accelerated from rest, gaining a kinetic energy k. Then

$$(k + m_0 c^2)^2 = E^2 = m_0^2 c^4 + p^2 c^2.$$

So

$$p = \frac{\sqrt{k}\sqrt{k + 2m_0 c^2}}{c}.$$

If the rest energy of the particle is much larger than the kinetic energy k, then we have approximately

$$p = \sqrt{k}\sqrt{2m_0},$$

which is equivalent to the classical formula with $p = m_0 v$,

$$k = \frac{m_0 v^2}{2} = \frac{p^2}{2m_0}.$$

Example The rest energy of the electron is

$$(9.11 \times 10^{-31})(3 \times 10^8)^2 = 8.2 \times 10^{-14}$$

Joules, which is 5.1×10^5 electron volts, or .51 mev. The rest energy of the proton is 935 mev, which is 1833 times bigger.

3 A Derivation of the Lorentz Transformation

This derivation is suggested by that in Bergmann **Basic Theories of Physics.** Peter Bergmann was one of Einstein's assistants at the Institute for Advanced Study at Princeton. Assume that the transformation is linear and preserves the metric

$$x_1^2 + x_2^2 + x_3^2 - c^2t^2.$$

Let the equations have the form

$$x_1' = \alpha x_1 + \beta t$$

$$x_2' = x_2$$

$$x_3' = x_3$$

$$t' = \gamma x_1 + \delta t$$

Let the primed frame have velocity v in the x_1 direction with respect to the unprimed frame. Then for any time t let $x_1 = vt$. Because this is the center of the primed frame, we have

$$x_1' = 0$$

so

$$0 = \alpha vt + \beta t = (\alpha v + \beta)t$$

Because t is arbitrary, we must have

$$\beta = -\alpha v$$

Now for any time t' let $x_1' = -vt'$. Because this is the coordinate of the center of the unprimed frame , which is moving relative to the primed frame with velocity $-v$, we must have $x_1 = 0$. From our assumed equations, we have

$$x_1' = -\alpha vt = -vt' = -v\delta t.$$

Thus

$$\delta = \alpha.$$

The transformation equations become

$$x_1' = \alpha(x_1 - vt)$$

$$x_2' = x_2$$

$$x_3' = x_3$$

$$t' = \gamma x_1 + \alpha t$$

Now we use the invariance of the quadratic form, which is a consequence of the constant nature of the velocity of light. We have

$$x_1^2 + x_2^2 + x_3^2 - c^2t^2 = x_1'^2 + x_2'^2 + x_3'^2 - c^2t'^2$$

$$= \alpha^2(x_1 - vt)^2$$

$$+x_2^2$$

$$+x_3^2$$

$$-c^2(\gamma x_1 + \alpha t)^2$$

$$= \alpha^2(x_1^2 - 2vtx_1 + v^2t^2)$$

$$+x_2^2$$

$$+x_3^2$$

$$-c^2(\gamma^2 x_1^2 + 2\gamma x_1 \alpha t + \alpha^2 t^2)$$

$$= x_1^2(\alpha^2 - c^2\gamma^2)$$

$$+x_1(-2vt\alpha^2 - c^2 2\gamma\alpha t)$$

$$+x_2^2$$

$$+x_3^2$$

$$-t^2c^2(\alpha^2 - \alpha^2\frac{v^2}{c^2}).$$

This must be a quadratic form, so the coefficient of x_1 must vanish. So

$$2\alpha t(v\alpha + c^2\gamma) = 0$$

Hence

$$\gamma = -\frac{v}{c^2}\alpha$$

In order to preserve the quadratic form, the coefficient of t^2c^2 must be 1, therefore

$$\alpha^2(1 - \frac{v^2}{c^2}) = 1$$

That is

$$\alpha = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Thus the Lorentz transformation equations are

$$x_1' = \alpha(x_1 - vt)$$

$$x_2' = x_2$$

$$x_3' = x_3$$

$$t' = \alpha(t - \frac{v}{c^2}x_1)$$

4 The Twins Paradox

Suppose one of a pair of twins travels off in a space ship at a relativistic speed so that his clock as measured on the earth runs 5 times slower than the clock of his twin, who remains on earth. He travels out for 5 years and then returns in another 5 years as measured on the earth. Then the twin has aged 10 years, but the returning twin will have only aged 2 years. It is often objected that the space traveling twin also sees the clock of the twin who remained on earth also running slow. But the problem is not symmetric in that the space traveling twin must accelerate to relativistic speeds and then reverse this acceleration to return and so on. So Einstein talked about this in another form in his original paper. He considered travel in a circle, approximated by a polygon, where velocity is uniform on each edge of the polygon. A clock would go off in a circular path and then return to its starting point. And would show less time elapsed than a stationary clock. This has been carried out in practice several times. A jet plane took an atomic clock in a path around the earth and returned, and indeed showed the time difference predicted by relativity. So in principle if you send your grandfather off in a spaceship, he could come home younger than you.

5 The Lorentz Boost

If we let $x_0 = ct, x_1 = x, x_2 = y$, and $x_3 = z$, then the Lorentz transformations become

$$x'_0 = \gamma(x_0 - \beta x_1)$$
$$x'_1 = \gamma(x_1 - \beta x_0)$$
$$x'_2 = x_2$$
$$x'_3 = x_3,$$

where

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

and

$$\beta = \frac{v}{c}.$$

Then the first two equations can be written as

$$x'_0 = \gamma x_0 - \gamma\beta x_1$$
$$x'_1 = \gamma x_1 - \gamma\beta x_0.$$

We have

$$\gamma^2 - (\gamma\beta)^2 = 1$$

We can take

$$\gamma = \cosh(\xi)$$

and

$$\gamma\beta = \sinh(\xi),$$

where ξ is called the Lorentz boost parameter. Then

$$x'_0 = \cosh(\xi)x_0 - \sinh(\xi)x_1$$
$$x'_1 = -\sinh(\xi)x_0 + \cosh(\xi)x_1,$$

which is similar to a rotation transformation, but with hyperbolic functions replacing trigonometric functions. See J. D. Jackson **Classical Electrodynamics**, 2nd edition, pp516-517. Also refer to the general treatment of the Lorentz group in S Sternberg **Group Theory and Physics** where it is shown that a general Lorentz transformation is the product of two rotations and a Lorentz boost.

6 Force And Acceleration

Newton's second law of motion is

$$\mathbf{F} = m\mathbf{a},$$

which is equivalent to force is equal the rate of change of momentum. The relativistic momentum is

$$\mathbf{P} = m\gamma\mathbf{v},$$

where by m we mean the constant rest mass, which we have written before as m_0. The relativistic version of Newton's law becomes

$$\mathbf{F} = \frac{d\mathbf{P}}{dt} = m\frac{d\gamma}{dt}\mathbf{v} + m\gamma\frac{d\mathbf{v}}{dt}.$$

$$= m\frac{d\gamma}{dt}\mathbf{v} + m\gamma\mathbf{a},$$

where $\mathbf{a}$ is the acceleration vector. Let $\mathbf{u}$ be a unit vector that is instantaneously parallel to the velocity $\mathbf{v}$.

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

Then we can resolve the acceleration into a sum of two instantaneous vectors, the first of which is parallel to $\mathbf{u}$ and the other perpendicular to it. Let

$$a_1 = \mathbf{a} \cdot \mathbf{u},$$

and let $\mathbf{w}$ be a unit vector in the direction of

$$\mathbf{a} - a_1\mathbf{u}.$$

then $\mathbf{w}$ is perpendicular to $\mathbf{u}$, and we may write

$$\mathbf{a} = a_1\mathbf{u} + a_2\mathbf{w},$$

where

$$a_2 = \mathbf{a} \cdot \mathbf{w}.$$

So a_1 and a_2 are the instantaneous components of the acceleration in the parallel and transverse directions to the velocity $\mathbf{v}$

Hence we can write our force equation as

$$\mathbf{F} = m\frac{d\gamma}{dt}v\mathbf{u} + m\gamma(a_1\mathbf{u} + a_2\mathbf{w})$$

$$= m(\frac{d\gamma}{dt}v + \gamma a_1)\mathbf{u} + m\gamma a_2\mathbf{w}.$$

We have

$$\frac{d\gamma}{dt} = (-1/2)\gamma^3(-1/c^2)\frac{dv^2}{dt}$$

$$= (-1/2)\gamma^3(-1/c^2)\frac{d\mathbf{v}\cdot\mathbf{v}}{dt}$$

$$= (-1/2)\gamma^3(-1/c^2)2\mathbf{v}\cdot\mathbf{a}$$

$$= (-1/2)\gamma^3(-1/c^2)2v\mathbf{u}\cdot\mathbf{a}$$

$$= \gamma^3(v/c^2)a_1.$$

Substituting this back into the force equation we have

$$\mathbf{F} = m(\frac{d\gamma}{dt}v + \gamma a_1)\mathbf{u} + m\gamma a_2\mathbf{w}$$

$$= m(\gamma^3(v^2/c^2)a_1 + \gamma a_1)\mathbf{u} + m\gamma a_2\mathbf{w}$$

$$= m\gamma a_1(\gamma^2(v^2/c^2) + 1)\mathbf{u} + m\gamma a_2\mathbf{w}$$

$$= m\gamma a_1(\frac{(v^2/c^2) + (1 - (v^2/c^2)}{1 - (v^2/c^2)}\mathbf{u} + m\gamma a_2\mathbf{w}$$

$$= m\gamma^3 a_1\mathbf{u} + m\gamma a_2\mathbf{w}.$$

Writing

$$\mathbf{F} = f_1\mathbf{u} + f_2\mathbf{w}$$

we see that the parallel component of the force is

$$f_1 = m\gamma^3 a_1,$$

and the transverse component is

$$f_2 = m\gamma a_2.$$

This contrasts with Newtonian mechanics where the force is proportional to the acceleration

$$\mathbf{F} = m\mathbf{a},$$

and the parallel and transverse equations would be

$$f_1 = ma_1,$$

and

$$f_2 = ma_2.$$

The relativistic equations must be taken into account in a particle accelerator when particles are traveling at relativistic speeds. Forces also transform so that forces are not necessarily the same in different coordinate systems. See Katz p59 for an example.

7 The Relativistic Form of Maxwell's Equations

Maxwell's Equations of electromagnetic theory are

$$\nabla \cdot \mathbf{D} = 4\pi\rho$$

$$\nabla \times \mathbf{H} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0$$

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

The Lorentz force on a charge is the sum of the electric and magnetic forces

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

In a certain sense Maxwell's equations imply relativity theory. There is a four dimensional covariant form of these equations. Because the divergence of a curl is zero, The magnetic induction **B** is given as the curl of a vector **A** called the vector potential. Also the electric field **E** is equal to the gradient of a scalar potential function Φ. These two potentials can be combined into a single 4-vector. Also the fields **E** and **B** can be combined into a 4 dimensional tensor called the Faraday field tensor **F**. Maxwell's four equations can be written as two covariant tensor equation. Covariant means that the form of the equation is preserved under a general Lorentz transformation, that is the form is the same in all inertial coordinate systems. See the chapter on special relativity in Jackson for an introduction to these ideas. The net result is that electricity and magnetism according to relativity are not really separate things. According to the relativistic formulation, it is possible for what appears to be an electric field in one coordinate system, to be a magnetic field in another inertial frame. See **The Feynmann Lectures on Physics** volume II, section 13.6, for the example of a current flowing in a wire and a charge outside of the wire moving parallel to the wire with velocity v. The charge then experiences a magnetic force toward the axis of the wire. However, with respect to a coordinate system moving with the charge, the charge seems to be at rest, and so have velocity zero. So it does not experience a magnetic force toward the center of the wire. Yet is must still experience a force toward the center of the wire. One can show that in this new coordinate system, the wire has a net electric charge, so the particle experiences an electric force. So there is just one electromagnetic force, and not two separate electric and magnetic forces. I once used this example to show that the conversion factor between the statcoulomb and the abcouloum is c the velocity of light (see my report: **A Mathematical Model and the Dynamic Simulation of an Electromagnetic Device**.)

8 4-Vectors

Relativity theory is not an isolated part of Physics. It permeates all of physics. We have seen that there is a 4-vector consisting of the three space dimensions and the time dimensions. There are other 4-vectors that transform in a manner similar to the space-time vector. Examples are a four potential consisting of the ordinary scalar electric potential and the 3 components of the vector potential, and the four potential consisting of the energy of a particle along with the three components of its momentum, and so on.

9 Classical Tensor Notation

Given two sets of coordinates x_1, x_2, x_3 and y_1, y_2, y_3, where each y_i, for $i = 1, 2, 3$ is a function of the x_1, x_2, x_3, and where each x_i, for $i = 1, 2, 3$

is a function of the y_1, y_2, y_3. We have the following relationship between differentials

$$dy_i = \sum_{j=1}^{3} \frac{\partial y_i}{\partial x_j} dx_j = \frac{\partial y_i}{\partial x_j} dx_j,$$

where we use the Einstein summation convention in the last expression. This convention assumes that repeated indices are summed, so that we avoid writing the summation symbol. On the other hand given a function f, consider the relationships between the sets of partial derivatives of f in the two coordinate systems.

$$\frac{\partial f}{\partial y_i} = \frac{\partial x_j}{\partial y_i} \frac{\partial f}{\partial x_j}.$$

These two transformation rules are different. The coefficients

$$\frac{\partial y_i}{\partial x_j},$$

occur in the first transformation rule, but the coefficients

$$\frac{\partial x_j}{\partial y_i}$$

occur in the second rule.

Suppose some entity has coefficients $C(i, x), i = 1, 3$ in the x coordinate system, and coefficients $C(i, y), i = 1, 3$ in the y. Suppose these coefficients transform linearly so that

$$C(i, y) = a(i, j)C(j, x).$$

Suppose

$$a(i, j) = \frac{\partial y_i}{\partial x_j},$$

then the coefficients of the entity transform as in the differential example above. In this case the coefficients are said to be the coefficients of a rank one contravariant tensor. In this case the coefficients are written with superscript indices

$$C(i, x) = C^i(x), C(i, y) = C^i(y).$$

And the transformation law is written as

$$C^i(y) = \frac{\partial y_i}{\partial x_j} C^j(x).$$

The convention is that indices to be summed should occur once as a superscript and once as a subscript.
Suppose on the other hand that

$$a(i,j) = \frac{\partial x_j}{\partial y_i},$$

then the coefficients of the entity transform as in the partial derivative example above. In this case the coefficients are said to be the coefficients of a rank one covariant tensor. In this case the coefficients are written with subscript indices

$$C(i,x) = C_i(x), C(i,y) = C_i(y).$$

And the transformation law is written as

$$C_i(y) = \frac{\partial x_j}{\partial y_i} C_j(x).$$

Notice that in transforming from the x coefficients to a y coefficients summing is done over an index j on x. So the cases

$$a(i,j) = \frac{\partial y_j}{\partial x_i},$$

and

$$a(i,j) = \frac{\partial x_i}{\partial y_j},$$

do not make sense.

The names covariant and contravariant correspond to transformations that, go respectively, like the "ordinary" transformation, and against the "ordinary" transformation. But what is the "ordinary" transformation. I am not aware of a good argument for designating the "ordinary" transformation. I think that the names covariant and contravariant should be taken just as a convention. These designations seem to be applied consistently in all the books that I am aware of.
Higher rank tensors are defined similarly. So suppose we have a rank 5 tensor $C^{ijk}_{lm}(x)$, which is contravariant of rank 3 and covariant of rank 2. Then the transformation rule is

$$C^{nop}_{qr}(y) = \frac{\partial y_n}{\partial x_i}\frac{\partial y_o}{\partial x_j}\frac{\partial y_p}{\partial x_k}\frac{\partial x_l}{\partial y_q}\frac{\partial x_m}{\partial y_r} C^{ijk}_{lm}(x).$$

So we know how to transform the tensor components, but what are the actual tensors? Tensors were thought of originally as the infinite set of all possible coefficients in all possible coordinate systems. This is not a very clear thought. It turns out that tensors can be interpreted as multilinear functionals defined on products of vector spaces. One such vector space is called the tangent space at a point. So this is something like the case of a 2d surface, say an ellipsoid, which at a point p has a tangent plane, which is a 2d vector space. This is for example where the velocity vectors of mechanics would live. The dual of this tangent space, in the linear algebra sense, is called the cotangent space. So contravariant and covariant tensors correspond to multilinear functionals on products of tangent and cotangent spaces.

10 Gauss's Intrinsic Geometry of Surfaces

While being involved in a project to measure the Earth, Gauss, who some would consider the greatest mathematician of all time, introduced the concept of intrinsic geometry of surfaces. Normally, when we examine a two dimensional surface we view it from outside in three space. But in setting up map projections and in doing earth measurements, one can not leave the surface of the earth, and hold it in the hand, so to speak, for examination. So Gauss thought about what properties can be determined, just by measurements on the surface itself. He came up with the idea of a general metric distance

$$ds^2 = \sum_{i=1}^{2}\sum_{j=1}^{2} g_{ij}dx^i dx^j,$$

that replaces the usual Euclidean or Pythagorean distance. The set of four numbers g_{ij} defines the metric tensor of the surface. For a general surface, these metric coefficients will vary at different points on the surface. This expression for ds^2 will turn out to define surface curvatures called the Gaussian curvature and the Mean curvature. Riemann extended these ideas to higher dimensional spaces and so invented what is now known as Riemannian Geometry. The fundamental idea of Einstein was the idea that these metric coefficient could be functions of local mass, producing a curvature of space, and explaining gravitation as a kind of force of acceleration due to the bending of space itself.

11 Tensor Analysis

Tensor analysis was invented by Tullio Levi-Civita during the end of the 19th century, and the beginning of the 20th century. The meaning of a tensor was originally rather vague. The coefficients of a tensor were certain multidimensional arrays of numbers with indices. These coefficients transformed according to certain rules under differential coordinate transformations. The tensor itself was defined as the collection of all such coefficient sets under all allowed transformations. Originally tensors were found to be quite useful and people learned how to manipulate them like crazy even though they did not completely understand what they were. Rigor finally came after the introduction of the concept of the differential manifold and related differential geometry. A differential manifold is a set of local coordinate maps from Euclidean space to the manifold object so that a differential manifold is a space that is locally Euclidean and such that neighborhood maps overlap in a smooth way. Thus the spherical earth is not Euclidean, but locally Euclidean. That is there are flat maps that represent the surface of the earth. But more than one such map is needed to cover the whole Earth. Located at a point on a differential manifold, is a tangent space where tangent vectors reside. In the case of a sphere, the tangent space at a point, is the plane tangent to the sphere at that point. However, the tangent space must be intrinsic. The manifold is not to be considered embedded in a higher space, so such tangent space must have some kind of abstract existence not depending on such an embedding. So anyway associated with each point of the manifold is a linear tangent space. On this tangent space (and on a dual cotangent space) one may define multilinear functionals, which are a slight generalization of linear maps to higher dimensions. These multilinear functionals are actually the tensors. This clears up what the conceptual meaning of a tensor is. But often one resorts to the old manipulation of coefficients for practical calculation with tensors. Tensors are often used in a pure Euclidean space, as in the concept of stress and strain tensors in elasticity theory. In the Euclidean space application, tensors are just an extension of matrices to higher dimensions. So anyway the original exposition of General Relativity by Einstein was in the language of tensors and Riemannian geometry. He had mathematical friends Michele Besso and Marcel Grossmann, who introduced him to this mathematics.

12 Differential Geometry

Differential geometry, differential manifold theory, and related concepts such as Lie Groups, Lie Algebras, Fibre Bundles, and so on, have become the language of physics. For example in String Theory, a string, representing a particle, is a one dimensional manifold embedded in a high dimensional space.

13 Riemannian Geometry

Various kinds of curvature tensors appear in relative theory to explain the gravitational force and the local geometry of space, as well as the geodesics, which are the curves of shortest length between two points, and give the path of light and photons in the nonEuclidean space-time.

14 Minkowski Space-Time

In 1908 Minkowski introduced the four dimensional space-time that gives equal status to coordinates of time and space. The metric defining the distance between two events is

$$(cdt)^2 - (dx_1^2 + dx_2^2 + dx_3^2)$$

and it is independent of the coordinate system (invariant). By multiplying the time by the imaginary number i , the the metric can be treated in some respects as an ordinary Euclidean 4 dimensional metric. But because this metric quadratic form is not positive definite, normal Riemannian geometry does not exactly apply, and so the geometry of relativity is not completely normal. Such a non-positive definite metric is called a non-Riemannian metric.

15 The Covariant Derivative and Parallel Transport

In Euclidian space with Cartesian basis functions, the covariant derivative of a vector field **Y** in the direction of a vector field **X** is the vector field of the directional derivatives of the components of $\mathbf{Y}$. If we use curvilinear coordinates, say spherical coordinates in this Euclidian space then in terms of the curvilinear basis vectors that change with position, the components are not just ordinary derivatives, but must be modified by the introduction of the Christoffel Symbols, which may be defined in terms of the metric coefficients g_{ij}, which define the distance element as

$$ds^2 = g_{ij} dq^i dq^j.$$

So they define the covariant derivative. If a curve and a tangent vector field along the curve say $\mathbf{T}$ have the property that the covariant derivative $\bar{D}_T T = 0$, then the curve is a geodesic, a curve of shortest distance in the space. A particle that does not experience a force moves along a geodesic. In a 3d Euclidian space this is uniform motion in a straight line. In the Newtonian theory of gravitation, mass produces a force which is action at a distance. So according to Mach, uniform motion is with respect to the center of mass of the universe. So when a particle moves it is being influenced by all the mass in the universe. Einstein thought this was a bit hard to believe. He wanted to replace this action at a distance by a local gravitational field. So the idea of general relativity is to modify Euclidean geometry by changing the metric coefficients so that the covariant derivative or connection gives the geodesics along which particles move. Einstein's famous elevator thought experiment resulted in his idea that there is an equivalence between gravitational force and acceleration.

In Newtonian mechanics an object that is not subject to a force continues in constant motion. If space is Euclidean, this just means that a velocity vector is a constant vector attached to the object and moving parallel to itself, that is translated along with the object. However, if space is non-Euclidean and warped, and the space is described locally by coordinate patches, what takes the place of a constant velocity vector when there is no force and no acceleration? To take an absurd example suppose the universe were only two dimensional and space was some sort of a wavy 2d surface. Then as an object moves with unaccelerated motion, the particle follows the surface, so its velocity vector located locally in some tangent space is transported continuously to some neighboring patch and tangent space. What constitutes this general transport of a vector that corresponds to parallel movement in Euclidean space? There is an entity called a connection on a manifold that allows us to transport velocity and acceleration vectors, so that velocity and acceleration are meaningful. Central to this idea is the concept of a covariant derivative.

16 The Riemann Curvature

If we parallel transport a vector on the surface of the earth along a spherical triangle made up of geodesics, we see that upon returning to the starting point, that the vector has rotated by some angle. So if we keep a vector parallel to itself, and move it along geodesics, which are the generalized straight lines, and we keep the vector always making the same angle to the straight line, then we find that the vector has rotated. This is because the surface is curved. We can transport a vector in such a way along an arbitrary closed curve by approximating the curve with short geodesics, in the same way as approximating a curve in Euclidean space with short line segments. We can define the Riemann curvature at a point as the limit of the angle change in parallel transport around a curve surrounding the point as the radius of the curve goes to zero. This works for 2d surfaces. Similarly Riemann curvature can be defined in higher dimensional Riemannian spaces. This curvature depends on the metric coefficients. In gravitational potential theory, the surface integral of the gravitational field is given by the amount of mass contained in the bounding surface. This is the analog of the Gauss's law in Electricity. In general we get Poisson's equation where the divergence of the field is proportional to mass density. This may be related to the Riemann curvature of space to give a semi-Riemannian metric defining the covariant derivative and hence the geodesics, i.e. the paths of particle motion.

One may also think of the directional derivative of a vector field Y with respect to a vector field X as a derivative in the direction of the field X at a point p. Meaning that if c is a curve in the direction of X at point P, then the directional derivative $\nabla_X Y$ is the derivative of Y along the curve, that is the derivative

$$\frac{dY(c(s))}{ds},$$

except that in computing the limit

$$\frac{Y'(c(s) - Y(c(0))}{ds}$$

$Y'(c(s))$ is the parallel translate of $Y(c(s))$ back to the tangent space at $c(0)$.

17 The Riemann Curvature Tensor

The Riemann curvature tensor is defined by (Hicks p59)

$$R(X,Y)Z = (\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - D_{[X,Y]}Z),$$

where [] is the bracket or commutator. It is a linear transformation that maps a tangent vector in the tangent space M_m into itself. Let its matrix components be R_{ij}. This measures the parallel translation. The components of a second Riemann curvature tensor is defined by

$$R^d_{abc} = R_{da}(X_b, X_c.$$

This has been constructed using differential operators, the commutator, or bracket. This curvature tensor is related to the Riemann curvature. Riemann curvature is related to the failure of a vector to return to its initial value when parallel transported around a loop. See Wald p37, and Hicks for a more mathematical treatment.

18 The Rici Tensor

The Rici tensor (changing notation) is the contraction of the Riemann tensor

$$R_{ac} = R^b_{abc}$$

The scalar curvature R is the trace of the Rici tensor

19 The Einstein Equation

To find curvature as influenced by mass-energy-momentum, (Einstein ,**the Meaning of Relativity**)we consider the source of the Newtonian gravitational field. Namely Poisson's equation for the gravitation potential ϕ (Einstein p82).

$$\nabla^2 \phi = 4\pi K \rho.$$

That is the divergence of the gradient of the potential ϕ, and is equal to a constant times the mass density. In the limit, the Relativistic theory must agree with the Newtonian theory. In relativity, the energy properties of matter are described as a stress-energy tensor T_{ab} where

$$T_{ab}v^a v^b,$$

corresponds to the mass density in classical Newtonian mechanics, and where v_a is the 4-velocity of the flow, a contribution to energy of both mass and motion. Einstein's field equation is

$$R_{ab} - \frac{1}{2}Rg_{ab} = 8\pi T_{ab}.$$

By solving this field equation one obtains the metric coefficients, which defines the covariant derivative, and thus the geodesics of the space, and so the directions of motion. Wald summarizes the entire content of general relativity by: **Spacetime is a manifold M on which there is defined a Lorentz metric g_{ab}. The curvature of g_{ab} is related to the matter distribution in spacetime by Einstein's equation.**

Solutions of problems in General Relativity involve solving difficult partial differential equations. They has been solved in closed form for only a few problems.

20 The Schwarzschild Solution

One of the first solutions to a general relativity equation was found by Schwarzschild. This problem was for the case of a static, spherically symmetric distribution of mass. Schwarzschild developed this solution, while serving in the first world war. He showed that if the density of mass of a body increases beyond a given point that the space of general relativity will bend so greatly as to generate a singularity at a point so that the geodesics (shortest path, the analog of the straight line), which light follows, curve towards this singularity so that all light and particles will flow into this singularity and never escape. Schwarzschild died in the war of illness probably related to his war service. Einstein in his book "The Meaning of Relativity" talks about the Schwarzschild solution. The metric for this solution is given in Wald (p124) as

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2 d\theta^2 + r^2 \sin^2(\theta)d\phi^2$$

As r goes to infinity, this metric goes to the Minkowski metric. Also the metric becomes singular at $r = 2M$.

$$r_s = 2M$$

is known as the Schwarzschild radius. For a body like the Sun, r_s is well inside the Sun, where the Schwarzschild solution is not valid.

Einstein in **The Meaning of Relativity,**, p 94, recommends the treatment of this solution as presented by Hermann Weyl in Raum-Zeit-Materie, English translation "Space Time Matter." Skimming through that book, it looks like difficult reading, with perhaps a little older notation. Also I did not find the name Schwarzschild mentioned. Herman Weyl was a very well known mathematician. His book Raum-Zeit-Materie is also supposed to contain the first introduction of the concept of the gauge transformation, which is a fundamental feature of modern particle physics. I think one should refer to Wald for a modern treatment of the Schwarzschild Solution. Lass also treats the Schwarzschild solution and its application to calculate the (pp328-334) orbit of Mercury, and the advance of the perihelion due to relativity.

21 Black Holes, The Big Bang, The Age of the Universe

These ideas are heavily dependent on the General Theory of Relativity.

22 Bibliography and References

[1] Arnold V I, **Mathematical Methods of Classical Mechanics**, 2nd edition, Springer-Verlag, Legendre Transformation, 1989

[2] Bergmann Peter G, **Introduction to the Theory of Relativity**. Dover reprint.

[3] Bergmann Peter G, **The Riddle of Gravitation**. Charles Scribner's Sons, New York,1968.

[4] Bergmann Peter **Basic Theories of Physics** 2 volumes, Dover.

[5] Einstein A., **Ist die Tragheit eines Korpers von sienem Energiegehalt abhangig?**, Annalen der Physik, 17, 1905. (Translation: Does the Inertia of a Body Depend Upon its Energy-Content?), This paper announced the equivalence of mass and energy.

[6] Einstein A., **The Meaning of Relativity**, 5th Edition, Princeton University Press, 1955. (The Stafford Little Lectures of Princeton University, May 1921)

[7] Einstein A., **Zur Electrodynamik bewegter Korper,** Annalen der Physik, 17,1905,, (Translation: "On the Electrodynamics of Moving Bodies"), This is the original Einstein paper on Special relativity.

[8] Einstein Albert, Lorentz H. A., Weyl Hermann, Minkowski H., **The Principal of Relativity,** A collection of original papers, notes by A. Sommerfeld, 1923, Dover reprint.

[9] Einstein Albert, **Relativity: the Special and General Theory**, 1931, Crown Publishers, New York.

[10] Eisberg Robert Martin, **Fundamentals of Modern Physics,** Revised Edition, John Wiley, 1967.

[11] Ellis George F. R., Williams Ruth M. **Flat and Curved Space-times,** Oxford University Press, 1990.

[12] Emery James D, **Differential Geometry**, Edition 2015, JDEU Press, (diffgeom.tex, diffgeom.pdf.)

[13] Emery James D, **Tensor Analysis in Euclidean Space**, Edition 2015, JDEU Press, (tensorana.tex, tensorana.pdf.)

[14] Feynman Richard P,Leighton Robert B, Sands Matthew, **The Feynman lectures on Physics**, Addison-Wesley, 3 Volumes, 1964.

[15] Firk Frank W. K., **Introduction to Relativistic Collisions**, The Henry Koerner Center for Emeritus Faculty, Yale University,
https://arxiv.org/ftp/arxiv/papers/1011/1011.1943.pdf

[16] Hicks Noel J, **Notes on Differential Geometry**, Van Nostrand, 1971.

[17] Jackson J. D. **Classical Electrodynamics**, 2nd edition, pp516-517.

[18] Katz, Robert **An Introduction to the Special Theory of Relativity,** Momentum, 1964.

[19] Krane Kenneth, **Modern Physics,** 2nd Edition, John Wiley, 1996.

[20] Lass Harry, **Vector and Tensor Analysis**, McGraw-Hill, 1950. (Section 141. Einstein's Law of Gravitation, Schwarzchild definition of metric coefficients, Advance of the perihelion in planetary motion.)

[21] Misner Charles W,Thorne Kip S, Wheeler John Archibald, **Gravitation**,W.H. Freeman and Company, 1973.

[22] Ohanian Hans C, Ruffini Remo **Gravitation and Spacetime,** 1994, Norton, 2nd ed.

[23] Pais Abraham, **Subtle is the Lord**.

[24] Panofsky Wolfgang K. H., Phillips Melba, **Classical Electricity and Mafnetism**, 2nd Edition, 1962, LHL QC518 .P35, Addison-Wesley.

[25] Prasanna A R, **Gravitation**. CRC Press, 2017, LHL QC178 .P74, (Relativity).

[26] Richtmyer F. K., Kennard E. H., Lauritsen T., **Introduction to Modern Physics,** 1955, McGraw-Hill.

[27] Riemann Bernhard **Über die Hypothesen welche der Geometrie zu Grunde liegen**, ("On the hypotheses which underlie geometry"), 1854. 1868 **On the hypotheses which lie at the foundation of geometry** translated by W.K.Clifford, Nature 8 1873 183, reprinted in Clifford's Collected Works.

[28] Schutz Bernard **A First Course in General Relativity** Cambridge

University Press, 1985.

[29] Schutz Bernard **Geometrical Methods of Mathematical Physics** Cambridge University Press, 1980.

[30] Stephani Hans **General Relativity: An Introduction to the theory of the gravitational field,** Cambridege University Press, 1982.

[31] Sternberg, S **Group Theory and Physics**

[32] Wald Robert M, **General Relativity**, University of Chicago Press, 1984, See chapter 6, the Schwarzschild Solution)

[33] Weyl Hermann, **Space, Time, Matter**, Dover, 1952.

23 Index