1 Introduction

1.1 What is relativity?

Physical theory, tested by experiment; mathematical structure derived from axioms; axioms must be valid in order for theory to be correct. Not relativism (which is philosophy or morality).

1.2 Speed of light

Finite: 17C by Ole Romer observing the moons of Jupiter. Many measurements since then; now *defined* to be 2.997792458×10^8 m/s. This is because length and time are now both measured using atoms and lasers.

1.3 Maxwell's theory

Mid 19C: light explained as an electromagnetic wave satisfying the following equations.

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

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

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

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right)$$

In free space, ρ and $\mathbf{J}$ zero, these lead to transverse waves with speed $c = 1/\sqrt{\epsilon_0 \mu_0}$. (Diagram)

1.4 Michelson-Morley experiment

Used laser interferometry to measure speed of light in orthogonal directions; rotation of the Earth means that many directions and velocities were measured. Light found to be constant. State of the art in 1880s, but now we can duplicate experiment to accuracy of 10^{-15}. (Diagram)

1.5 Lorentz contraction and time dilation

Lorentz proposed a solution: moving objects suffered a contraction along their length, and a slowing of their time.

1.6 Einstein's axioms (1905)

Relativity: physical laws are valid in all inertial frames of reference. The speed of light is the same in all inertial frames of reference. (More on this later)

1.7 Poincare space-time

Poincare showed that Einstein's theory was most simply expressed as a 4D space-time with indefinite metric $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$.

1.8 General relativity

Einstein included gravity (and the Newton's constant G) in 1917. Special relativity valid as long as GM/rc^2 small.

1.9 Quantum field theory

Quantum mechanics introduces another constant $\hbar$. Special relativistic dynamics valid for energy-momentum conservation of particles but needs full field theory for calculation of scattering proabilities etc.

1.10 Quantum gravity

No complete theory exists for the regime in which c, G and $\hbar$ are all important - experiments all but impossible, and it has shown too hard for the theorists so far (Einstein used only a minimum of experiment).

1.11 Dimensional analysis

We will see that c plays an important role in special relativity, but in a sense just converts between units of length and time. Often in the literature people use units where $c = 1$: this means expressing equations in terms of the dimensionless v/c, times as ct become lengths, and masses mc^2 and momenta pc become energies. We will usually keep the c for clarity.

1.12 Current status

Where it is valid, special relativity has been confirmed by all manner of experiments on high energy particles in accelerators and from space, and precision experiments involving very accurate clocks. The opposition was quite substantial until the 1950s - either the experiments improved or the older generation of physicists died out!

2 The Lorentz transformations

2.1 Galilean relativity

Einstein's first axiom is borrowed from the classical physics of Galileo and Newton. If we consider two observers in the "standard configuration", we write

$$x' = x - vt$$

$$y' = y$$
$$z' = z$$

then we differentiate to get

$$u'_x = u_x - v$$
$$u'_y = u_y$$
$$u'_z = u_z$$

and finally

$$\mathbf{a}' = \mathbf{a}$$

2.2 Invariance of Newtonian mechanics

So both observers see the same force, Newton's second law is invariant. What about conservation of momentum and energy?

$$\mathbf{p'}_f - \mathbf{p'}_i = \sum_j (m_j(\mathbf{u'}_{jf} - \mathbf{u'}_{ji})) = \sum_j (m_j(\mathbf{u}_{jf} - \mathbf{u}_{ji})) = \mathbf{p}_f - \mathbf{p}_i = 0$$

$$E'_f - E'_i = \frac{1}{2}\sum_j (m_j(u'^2_{jf} - u'^2_{ji})) = \frac{1}{2}\sum_j (m_j(u^2_{jf} - u^2_{ji} - 2\mathbf{v}\cdot(\mathbf{u}_{jf} - \mathbf{u}_{ji}))$$
$$= E_f - E_i - 2\mathbf{v}\cdot(\mathbf{p}_f - \mathbf{p}_i) = 0$$

Energy and momentum are different to different observers, but they are still both conserved.

2.3 Noninvariance of Maxwell's equations

Under Galilean transformations, velocities get shifted; thus Maxwell's equations (containing c) cannot be invariant. For sound waves we understand that the equations hold only for an observer at rest with respect to the medium, but what is the medium for light? In the 19C, such a medium was postulated, called the luminiferous aether.

2.4 Michelson-Morley experiment

The aim of the experiment was to determine the velocity of the Earth relative to the aether. It did this by measuring the speed of light in two orthogonal directions. (Diagram). Even though the Earth moves with different velocities at different times in its orbit, no variation in the speed of light was observed. Conclusion: Speed of light is the same to every observer!

2.5 Loss of absolute simultaneity

Since this is not compatible with Galilean relativity, we must change the axioms: in this case, the unwritten "obvious" transformation

$$t' = t$$

Thus, two events which occur simultaneously for one observer may not for an observer moving with respect to the first. Some observers consider A before B, while others consider B before A. NB: As we will see, there is no loss of causality, since a signal from A to B or B to A must travel faster than light, which is prohibited in special relativity.

2.6 Frames of reference

We must now carefully define what we mean by an observer's "frame of reference". An *event* is a particular point at a particular time - different observers will give the same event different positions and times. Distance is defined using standard rods and time using standard clocks, both at rest with respect to an observer S. Each observer has a corresponding set of rods and clocks. Two events A and B are defined to be simultaneous in a particular reference frame if light beams from A and B meet at a point equidistant between A and B according to the rods. The clocks at A and B are set to the same time. NB1:Note that we are making simple physical assumptions, for example, if A and B are simultaneous while B and C are simultaneous, then we must have that A and C are also simultaneous. This holds automatically if space is Euclidean, clocks all run at the same rate, and light always travels at the same speed. The first two are broken in general relativity, and as a consequence there is no unique definition of simultaneity. NB2: the observer doesn't actually "see" an event A at time t_A; only after the light has had time to reach the observer.

2.7 The Lorentz transformations

Now we are going to derive the equations corresponding to the Galilean transformations using Einstein's two axioms. The observers are defined to be in "inertial reference frames", for which Newton's first law holds. Since all linear motion is mapped to linear motion, the transformations themselves must be linear. Let τ be the time registered by a freely moving clock, then ($x^\mu = ct, x, y, z$; include sums)

$$\frac{dx'^\mu}{d\tau} = \sum_\nu \frac{\partial x'^\mu}{\partial x^\nu}\frac{dx^\nu}{d\tau}$$

$$\frac{d^2x'^\mu}{d\tau} = \sum_{\nu\sigma} \frac{\partial^2 x'^\mu}{\partial x^\nu x^\sigma}\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau} + \sum_n u\frac{\partial x'^\mu}{\partial x^\nu}\frac{d^2x^\nu}{d\tau^2}$$

Since the cross derivative must be zero, and the origins of the two reference frames are aligned, we must have

$$x^{\mu'} = \sum_n u \Lambda^{\mu'}_{\nu} x^{\nu}$$

Now we must have $y' = y$ and $z' = z$ since any argument for $y' > y$ would need to give the opposite when $v \to -v$, ie exchanging S and S'. Now the known relative velocities of S and S' give

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

$$t' = v\eta(v)x + \alpha(v)t$$

Use reflection symmetry again ($x \to -x$, $x' \to -x'$ and $v \to -v$) we find that γ, η and α are even functions. Exchanging S and S':

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

$$t = -v\eta x' + \alpha t'$$

substituting and equating coefficients we find

$$\eta = \frac{1 - \gamma^2}{\gamma v^2}$$

$$\alpha = \gamma$$

If we write $\gamma = 1$ at this point we are back to the Galilean transformations. Now we invoke the uniformity of the speed of light c. Let a beam of light issue from the origin along the x axis, then we have $c = x/t = x'/t'$. Solving we find

$$c = \frac{x'}{t'} = \frac{\gamma(c - v)}{\frac{1-\gamma^2}{\gamma v} c + \gamma} = \frac{\gamma^2 v(c - v)}{(1 - \gamma^2)c + \gamma^2 v}$$

$$c^2 = \gamma^2(cv - v^2 + c^2 - cv)$$

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

so $\gamma = 1/\sqrt{1 - v^2/c^2}$ and hence

$$x' = (x - vt)/\sqrt{1 - v^2/c^2}$$

$$t' = (t - vx/c^2)/\sqrt{1 - v^2/c^2}$$

are the final form of the Lorentz transformations.

3 Kinematics

3.1 How big is γ?

Car at 60 miles an hour, $\gamma = 1 + 5 \times 10^{-15}$
Escape velocity of Earth (7 miles a second), $\gamma = 1 + 5 \times 10^{-10}$
Outer electron in atom ($v/c = 1/137$), $\gamma = 1 + 3 \times 10^{-5}$
Half the speed of light, $\gamma = 1.1547$
Electron in particle accelerator, $\gamma = 10^5$ Cosmic ray proton $\gamma = 10^{10}$

3.2 Length contraction

Consider a rod of length L at rest in S'. Its ends are described by the equations $x' = 0$, $x' = L$. In S, these become

$$x = vt$$

$$x = vt + L/\gamma$$

that is, moving objects are apparently contracted. Of course both observers consider that objects in the other's frame are contracted; this is possible because the length is measured at fixed t in frame S and at fixed t' in frame S'.
Garage paradox: suppose you have a garage of length L and you want to fit a car of length $2L$ in by moving it at $c\sqrt{3}/2$. In the frame of the car, the garage has length $L/2$, so the car hits the back wall before it is fully inside. However, the back of the car does not know (via light signal) that it has collided until after it is fully inside, so both observers agree that the rear end of the car enters the garage. In full, the front of the garage takes a time $(3L/2)/(c\sqrt{3}/2) = \sqrt{3}L/c$ to reach the back of the car, but the impact takes $2L/c$. This illustrates another important point, that it is impossible for an extended object (such as the car) to be rigid. (Spacetime diagrams)

3.3 Time dilation

Consider a clock at rest in S', with $x' = 0$. The Lorentz transformations give $x = vt$, hence

$$t' = \gamma(t - v^2 t/c^2) = t/\gamma$$

The moving clock appears to slow down by a factor γ. The time dilation effect has been amply verified in experiments involving the decay of fast moving particles.
The twin paradox: Suppose A remains on Earth, while B travels at velocity v for some time before returning to Earth. B would have aged by a factor γ less than A. How is this possible since the Lorentz transformations are symmetrical? Remember that they hold only for *inertial* frames of reference. A is in an inertial frame, but B is not. An inertial frame moving with B in the outward phase of the journey would lead to a stronger time dilation effect in the return journey, enough to reach the same conclusion. (Spacetime diagrams). Circular motion is also non-inertial.

3.4 Velocity transformation

We can generally relate velocities and accelerations in S' to S by differentiating the Lorentz transformations.

$$u'_x = \frac{dx'}{dt'} = \frac{dx'/dt}{dt'/dt} = \frac{u_x - v}{1 - u_x v/c^2}$$

$$u'_y = \frac{u_y}{\gamma(1 - u_x v/c^2)}$$

$$u'_z = \frac{u_z}{\gamma(1 - u_x v/c^2)}$$

Velocity addition is obtained by changing the signs.
Example: $u = c$, then $u' = c$ as we already know.
Example: $u = v$, then $u' = 2v/(1 + v^2/c^2)$. For the twin paradox, we have that A experiences a time $2L/v$ and B a time $(2L/v)\sqrt{1 - v^2/c^2}$ by looking at A's reference frame. In B's (outgoing) reference frame, we have

$$x_A = -vt$$

$$x_B = -(2v/(1 + v^2/c^2))(t - (L/v)\sqrt{1 - v^2/c^2})$$

Solving these equations gives

$$t(\frac{2v}{1 + v^2/c^2} - v) = \frac{2L}{1 + v^2/c^2}\sqrt{1 - v^2/c^2}$$

$$t = \frac{\frac{2L}{1+v^2/c^2}\sqrt{1 - v^2/c^2}}{v(\frac{1-v^2/c^2}{1+v^2/c^2}})$$

$$t = (2L/(v\sqrt{1 - v^2/c^2}))$$

and hence A's time of $2L/v$. For B's time we compute the γ factor for the new doubled velocity as $(c^2 + v^2)/(c^2 - v^2)$, and finally the same time as given above (skip algebra!).

3.5 Acceleration

Differentiating the x-eq once again, we get

$$a'_x = \frac{a_x}{\gamma^3(1 - u_x v/c^2)^3}$$

If we set $u_x = v$ we find the acceleration as measured in the particle's own rest frame, *proper acceleration* which is simply

$$\alpha = a'_x = a\gamma^3 = \frac{d}{dt}(\gamma u_x)$$

Note that acceleration is no longer an invariant, as in the Galilean transformations. It will turn out when we study dynamics that the most appropriate generalisation of momentum involves this γu.
Example: Consider a particle moving with constant proper acceleration. We immediately integrate to find

$$\alpha t = \gamma u$$

Solve for u:

$$u = \frac{\alpha c t}{\sqrt{c^2 + \alpha^2 t^2}}$$

and integrate again...

$$x = \frac{c}{\alpha}\sqrt{c^2 + \alpha^2 t^2}$$

which is the equation of a hyperbola. Note that $x > ct$ so a photon chasing the particle from the origin never catches up.

4 Optics

4.1 Speed of light in medium

In a medium of refractive index $n > 1$, the speed of light is known to be $u' = c/n$ relative to the medium. In terms of Maxwell's equations, $n = \sqrt{\epsilon\mu/\epsilon_0\mu_0}$. For example: water has $n = 1.33$ and glass has $n = 1.5 - 1.6$. Thus for small velocities v of the medium we have

$$u = \frac{u' + v}{1 + u'v/c^2} \approx (u' + v)(1 - u'v/c^2) \approx u' + v(1 - u'^2/c^2) = u' + v(1 - 1/n^2)$$

so the greater the refractive index, the more the light is "dragged" by the medium.

4.2 The Doppler effect

Suppose a source of waves moves with speed u at an angle θ to the x-axis. The time interval between successive crests is $\Delta t'$ in the frame of the source. In the observer's frame, these are time dilated to

$$\Delta t_s = \gamma \Delta t'$$

However, they must also travel an extra distance $u \cos\theta \Delta t_s$, so the total time between crests as observed by the observer is

$$\Delta t = \gamma \Delta t'(1 + u\cos\theta/c)$$

or in frequency,

$$f'/f = \gamma(1 + u\cos\theta/c)$$

In the nonrelativistic limit, we have no time dilation, so $\gamma = 1$. In relativity, the time dilation causes a transverse Doppler effect, of order v^2/c^2. The special case of $\theta = 0$ leads to a formula for a source moving directly away:

$$f'/f = \sqrt{\frac{c+v}{c-v}}$$

and $\theta = \pi$ for directly towards:

$$f'/f = \sqrt{\frac{c-v}{c+v}}$$

Note that in the astronomy literature, the Doppler redshift is measured in terms of z, where

$$1 + z = f'/f$$

4.3 Aberration

Consider a light ray coming from a direction α with respect to the x-axis. Then its velocity is $(-c\cos\alpha, -c\sin\alpha)$ in the S frame. Using the velocity transformation formula,

$$\cos\alpha' = \frac{\cos\alpha + v/c}{1 + (v/c)\cos\alpha}$$

For the motion of the Earth around the sun, $v = 30km/s$ so $v/c = 10^{-4}$, and the maximum deviation is about 10^{-4} radians equals 20 seconds of arc.

5 Spacetime

5.1 The interval

Let us consider the proper time of an arbitrarily moving clock. The time read by such a clock will be given by an integral over the appropriate *gamma*-factor:

$$c\tau = \int cdt/\gamma = \int cdt\sqrt{1 - v^2/c^2} = \int \sqrt{c^2dt^2 - dx^2 - dy^2 - dz^2} = \int ds$$

It is easily verified that the quantity ds^2 is left invariant under Lorentz transformations, and so generalises the concept of distance to four dimensional spacetime:

$$\begin{aligned} c^2dt'^2 - dx'^2 &= \gamma^2[c^2dt^2 - 2vdtdx + v^2dx^2/c^2 - dx^2 + 2vdxdt - v^2dt^2] \\ &= \gamma^2(1 - v^2/c^2)[c^2dt^2 - dx^2] \\ &= c^2dt^2 - dx^2 \end{aligned}$$

We can write this in tensor notation as

$$ds^2 = \sum_{\mu\nu} g_{\mu\nu}dx^\mu dx^\nu$$

with $g_{\mu\nu}$ a matrix with diagonal elements $(1, -1, -1, -1)$. Note that other conventions are in use: "Spacelike", $(-1, 1, 1, 1)$ and (rarely now) "Euclidean", $(1, 1, 1, 1)$ with $x^0 = ict$. In our timelike convention, we have that timelike intervals have $ds^2 > 0$, lightlike intervals have $ds^2 = 0$ and spacelike intervals have $ds^2 < 0$.

5.2 2D Euclidean analogy

Flatlanders move mostly in one direction t measured in feet, and only slightly in the orthogonal direction x measured in nanometres. Their deviation from the t direction is $v = x/t = c \tan\theta$ where c is a fundamental constant equal to 3×10^8 nanometres per foot. To transform from one frame to the other, they rotate:

$$x' = \cos\theta x - \sin\theta ct$$

$$t' = \sin\theta x/c + \cos\theta t$$

To write this in terms of velocity we use

$$\cos\theta = \frac{1}{\sqrt{1 + \tan^2\theta}}$$

$$\sin\theta = \frac{\tan\theta}{\sqrt{1 + \tan^2\theta}}$$

to obtain

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

$$t' = \gamma(t + vx/c^2)$$

$$\gamma = 1/\sqrt{1 + v^2/c^2}$$

which, apart from some signs are just the Lorentz transformations! The velocity addition formula is

$$c \tan\theta_1 + \theta_2 = \frac{c(\tan\theta_1 + \tan\theta_2)}{1 - \tan\theta_1 \tan\theta_2} = \frac{v_1 + v_2}{1 - v_1 v_2/c^2}$$

again the same but with different signs.

5.3 Rapidity

What is our analog of θ? To make the correct signs, we need hyperbolic rather than circular functions. If we introduce $v = c \tanh\phi$ we obtain

$$x' = \cosh\phi x - \sinh\phi t$$

$$t' = -\sinh\phi x + \cosh\phi t$$

$$\gamma = \cosh\phi$$

and the velocity addition formula is just the addition formula for tanh. The parameter ϕ is called the rapidity, and is additive (rather than velocity), at least in one dimension. It is often handy to use the analytic form for the inverse tanh function:

$$\phi = \frac{1}{2} \ln \frac{c+v}{c-v}$$

5.4 The Lorentz group

It is clear that the set of linear transformations that preserve the interval form a group, called the Lorentz group. The matrices have determinant 1:

$$ds^2 = dx^T g dx = dx'^T g dx' = dx^T \Lambda^T g \Lambda dx$$

This is true for all x, thus we have

$$g = \Lambda^T g \Lambda$$

Taking the determinant of both sides we find:

$$det\Lambda = \pm 1$$

This group is generated by two sets of continuous transformations: rotations and boosts, and two discrete symmetries, spatial reflection and time reversal. Spatial reflection is not continuous since it corresponds to determinant -1 rather than 1, and time reversal is not continuous since the interval is negative in the intervening spacelike region. Transformations that preserve the spatial orientation are called *proper*, and those that preserve the time orientation are called *orthochronos*. Note that in general, the rotations and boosts do not commute. If translations are also included, the group is called the inhomogeneous Lorentz group, or the Poincare group.

5.5 scalars and vectors

We can now classify various quantities according to their transformation properties under the Lorentz group. In doing so, we will be able to write equations which are obviously invariant under the Lorentz transformations. The simplest quantities, *scalars* are invariant, that is, they have the same values in all inertial frames. Examples are the (rest-)mass of a particle, the temperature, the interval Δs^2. Then we have vectors with contravariant (upper) indices, which transform as space-time intervals:

$$\Delta x^{\mu'} = \sum_\nu \Lambda^{\mu'}_{\ \nu} \Delta x^\nu$$

Recall that $\mu = 0, 1, 2, 3$ corresponding to (ct, x, y, z).
Note that in general relativity, the space-time intervals must be infinitesimal dx^μ, and Λ is the matrix of derivatives of arbitrary differentiable coordinate transformations.

It is easy to show that the set of such quantities forms a vector space $\mathcal{V}$ of dimension 4:

$$Z^\mu = X^\mu + Y^\mu$$

is a vector,

$$W^\mu = aX^\mu$$

is a vector if a is a scalar,

$$a(X^\mu + Y^\mu) = aX^\mu + aY^\mu$$

The zero vector has components $(0, 0, 0, 0)$. Every vector can be constructed as a linear combination of the four basis vectors $E_0 = (1, 0, 0, 0)$, $E_1 = (0, 1, 0, 0)$, $E_2 = (0, 0, 1, 0)$, $E_3 = (0, 0, 0, 1)$ in a particular reference frame. The examples of vectors that we have so far are positions X^μ, finite intervals ΔX^μ and infinitesimal intervals dx^μ. Note that only the infinitesimal intervals are vectors in GR, because they are required to be invariant under arbitrary differentiable transformations, not necessarily linear.

5.6 1-forms (dual vectors, co-vectors)

Recall the construction of the dual vector space $\mathcal{V}^*$: we consider all possible linear functionals of vectors $\rho : \mathcal{V} \to \mathbf{R}$, that is,

$$\rho(aX + bY) = a\rho(X) + b\rho(Y)$$

The most general linear functional of a vector X is of the form

$$\rho(X) = \sum_\mu \rho_\mu x^\mu$$

note that X is also a linear functional of the covariant vector (or "1-form") ρ. There is a dual basis for one-forms $\{\omega^\alpha\}$, satisfying

$$\omega^\alpha(E_\beta) = \delta^\alpha_\beta$$

Here the Kronecker delta is one if $\alpha = \beta$, zero otherwise. This duality is the basis for the Einstein summation convention, where we always sum over repeated upper and lower indices. Note, this convention can be applied to any finite dimensional vector space - for a complex space, dotted and undotted indices are used to signify complex conjugation; for an infinite dimensional space, we run into problems with the order of infinite sums or integrals.
How do the covariant components transform? $\omega(X)$ is a scalar, so it is unaffected. So we have

$$\omega(X) = \omega_{\mu'} X^{\mu'} = \omega_{\mu'} \Lambda^{\mu'}_{\ \nu} X^\nu$$

for any X. Thus we need

$$\omega_{\mu'} = \Lambda^{\sigma}_{\ \mu'} \omega_\sigma$$

with

$$\Lambda^{\sigma}_{\mu'}\Lambda^{\mu'}_{\nu} = \delta^{\sigma}_{\nu}$$

that is, the two transformations are inverses. An example of a one-form is the derivative of a scalar,

$$\rho_\alpha = \frac{\partial}{\partial x^\alpha}\phi = \phi_{,\alpha}$$

$$\rho_{\alpha'} = \frac{\partial}{\partial x^{\alpha'}}\phi = \frac{\partial x^\beta}{\partial x^{\alpha'}}\frac{\partial}{\partial x^\beta}\phi$$

5.7 Tensors

We introduce tensors as linear functionals of many vectors and one-forms. For example, given vectors X, Y and 1-forms σ, ρ we have a fourth rank tensor

$$A(\sigma, X, Y, \rho) = A^{\alpha}{}_{\beta\gamma}{}^{\delta}\sigma_\alpha X^\beta Y^\gamma \rho_\delta$$

which transforms as

$$A^{\alpha'}{}_{\beta'\gamma'}{}^{\delta'} = \Lambda^{\alpha'}_{\alpha}\Lambda^{\beta}_{\beta'}\Lambda^{\gamma}_{\gamma'}\Lambda^{\delta'}_{\delta}A^{\alpha}{}_{\beta\gamma}{}^{\delta}$$

New tensors from old! We can add tensors of the same positioning of indices.

$$C(\sigma, X, Y, \rho) = A(\sigma, X, Y, \rho) + B(\sigma, X, Y, \rho)$$

We can multiply by a scalar.

$$B(\sigma, X, Y, \rho) = aA(\sigma, X, Y, \rho)$$

We can contract upper and lower indices.

$$D(X, \rho) = \sum_\alpha C(\omega^\alpha, X, E_\alpha, \rho)$$

We can construct an outer product.

$$C(\sigma, X, Y, \rho) = A(\sigma, X, Y)B(\rho)$$

Note that if C and A are tensors in such an expression, then B must be also (tensor "division"). We can rearrange the arguments of like type (generalised transpose).

$$B(X, Y, Z) = A(Y, Z, X)$$

We can differentiate a tensor.

$$C(\sigma, X, Y) = (\partial B(\sigma, X))(Y)$$

In general, tensors with more than one upper and/or lower index can have a variety of symmetries; for the case of two indicies we have a decomposition into symmetric and antisymmetric parts:

$$T^{\alpha\beta} = T^{(\alpha\beta)} + T^{[\alpha\beta]}$$

$$T^{(\alpha\beta)} = \frac{1}{2}(T^{\alpha\beta} + T^{\beta\alpha})$$
$$T^{[\alpha\beta]} = \frac{1}{2}(T^{\alpha\beta} - T^{\beta\alpha})$$

Example: Suppose we have an expression

$$(A^{\mu\nu} - A^{\nu\mu})B^{\alpha}_{\ ,\nu\mu}$$

This must be zero: we can exchange the order of the derivatives, relabel ν and μ, and hence show that this expression is equal to its negative. In general, contraction of symmetric and antisymmetric indices is zero.

5.8 The metric

So far we have not introduced the concept of a metric. In the case of special relativity, we know that the interval is a scalar,

$$ds^2 = g_{\mu\nu}dx^{\mu}dx^{\nu}$$

so $g_{\mu\nu} = diag(1, -1, -1, -1)$ must be a symmetric second rank tensor. Thus we can define an inner product of vectors,

$$A \cdot B = g_{\mu\nu}A^{\mu}B^{\nu} = g(A, B)$$

which must be a scalar, symmetric in A and B, and distribute addition. We can define orthogonality by $A \cdot B = 0$ and the square of a vector (can be negative) by $A \cdot A$. A vector is timelike if this is positive, lightlike if zero, and spacelike if negative. We can also use the metric to raise and lower indices

$$A_{\mu} = g_{\mu\nu}A^{\nu}$$
$$A^{\nu} = g^{\nu\mu}A_{\mu}$$

where $g^{\nu\mu}$ is the matrix inverse of $g_{\mu\nu}$ (numerically the same matrix). Therefore, the lower components of a vector are $(A^0, -A^1, -A^2, -A^3)$. Since the metric tensor is constant, the raising and lowering commutes with differentiation:

$$A_{\mu,\nu} = (g^{\mu\rho}A^{\rho})_{,\nu} = g_{\mu\rho}A^{\rho}_{\ \nu}$$

Note: In general relativity, the metric tensor is not constant; however the above relation is maintained by modifying the derivative.

5.9 3-vectors

The rotation group is a subgroup of the Lorentz group, and the spatial components of a 4-vector are a 3-vector under this subgroup. We will denote 3-indices using Latin letters a, b, c. If we have a 4-vector $A = (A^0, \mathbf{a})$ and $B = (B^0, \mathbf{b})$ then the inner product is

$$A \cdot B = A^0B^0 - \mathbf{a} \cdot \mathbf{b}$$

where $\mathbf{a} \cdot \mathbf{b}$ is the usual three dimensional inner product. Note that if, as in the usual 3D case, the metric components are all one, we don't need upper and lower indices, as both components are always the same.

6 Particle mechanics

6.1 Proper time, four velocity and acceleration

Now we have the 4-tensor formalism, we can construct equations which we know will transform correctly under the Lorentz transformations. We began our discussion of the interval using the example of proper time,

$$d\tau^2 = ds^2/c^2 = g_{\mu\nu} dx^\mu dx^\nu$$

Derivatives with respect to the proper time of a particle are evaluated simply as

$$\frac{d}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt} = \gamma \frac{d}{dt}$$

We construct a Lorentz invariant 4-velocity by differentiating the position 4-vector with respect to this proper time,

$$U^\mu = \frac{d}{d\tau} x^\mu = (\gamma c, \gamma \mathbf{u})$$

Note that in the reference frame of the particle, the 4-velocity becomes $(c, 0, 0, 0)$. Thus $U \cdot F$ gives the component cF^0 as measured in the reference frame of the particle for some 4-vector F. We also have $U^2 = U \cdot U = c^2$, easiest to prove in the reference frame of the particle, but true in all reference frames as it must be. This also shows that U is a timelike, future pointing 4-vector. If we have two particles with 4-velocities U and V, their components in V's reference frame are $U = (\gamma c, \gamma \mathbf{u})$, $V = (c, 0)$, thus

$$U \cdot V = c^2 \gamma$$

where the γ is for the relative velocity of the particles.

Let us rederive the transformation law for velocity. We have

$$U^{\alpha'} = \Lambda^{\alpha'}_{\beta} U^\beta$$

Thus for the t and x components, we have

$$\begin{pmatrix} \gamma' c \\ \gamma' u' \end{pmatrix} = \begin{pmatrix} \Gamma & -\Gamma v/c \\ -\Gamma v/c & \Gamma \end{pmatrix} \begin{pmatrix} \gamma c \\ \gamma u \end{pmatrix}$$

Thus

$$\gamma' c = \Gamma \gamma c - \Gamma \gamma u v/c$$
$$\gamma' u' = -\Gamma \gamma v + \Gamma \gamma u$$

Dividing, we find

$$u'/c = \frac{u - v}{c - uv/c}$$

as expected.

We take another derivative with respect to proper time,

$$A^\mu = \frac{d}{d\tau}U^\mu = (c\gamma\frac{d\gamma}{dt}, \mathbf{u}\gamma\frac{d\gamma}{dt} + \gamma^2\mathbf{a}$$

In the instantaneous rest frame of the particle, this is simply $(0, \mathbf{a})$. Thus the proper acceleration can be written

$$\alpha^2 = -A^2$$

and A is spacelike or zero. Also, since U^2 is a constant, the derivative of this leads to

$$U \cdot A = 0$$

so the 4-velocity and 4-acceleration are orthogonal 4-vectors.

Example: uniform circular motion. We have

$$x = (ct, r\cos\omega t, r\sin\omega t, 0)$$

Differentiating, we find

$$U = \gamma(c, -u\sin\omega t, u\cos\omega t, 0)$$

with $u = r\omega$. Differentiating again, we have

$$A = \gamma^2(0, -a\cos\omega t, -a\sin\omega t, 0)$$

with $a = u\omega = r\omega^2$ and u and hence γ do not depend in time. Hence

$$\alpha = \sqrt{-A^2} = \gamma^2 a = \frac{r\omega^2}{1 - r^2\omega^2/c^2}$$

6.2 Energy and momentum

In order to talk about dynamics, we need more than the geometry of spacetime. We need to ask what is observed in experiments on fast moving particles, and compare the results with laws in terms of 4-vectors and tensors. We begin with conservation of energy and momentum, which are valid even when we know nothing about the details of the forces (as in collisions of elementary particles). We replace the nonrelativistic momentum $\mathbf{p} = m\mathbf{u}$ by its natural analog

$$P^\mu = mU^\mu = (mc\gamma, m\gamma\mathbf{u}) = (E/c, \mathbf{p})$$

$$P^2 = m^2U^2 = m^2c^2$$

Here, m is the mass of the particle as measured in its rest frame ("rest mass"). Note that some authors (notably Rindler) use m_0 for rest mass and a symbol $m = \gamma m_0$ to indicate a mass that increases with velocity. P^μ is the 4-momentum, and the claim is that E and $\mathbf{p}$ reduce to energy and 3-momentum in the nonrelativistic limit. First, it is clear that $\mathbf{p}$ is just the old momentum when $\gamma = 1$. If we expand the energy in powers of v/c we find

$$E = \gamma mc^2 = mc^2 + mu^2/2 + 3mu^4/8c^2 + \dots$$

The first term is a constant, and so would be ignored in nonrelativistic calculations. It is, however, important if we want to create or destroy particles (allowed under certain conditions). It is called “rest energy”. The second term is the usual kinetic energy, and the following terms are corrections which are small when the velocity is small. Kinetic energy is the full energy minus the rest energy

$$K = E - mc^2 = (\gamma - 1)mc^2$$

If the rest energies are not equal before and after an interaction, the difference is accounted for in kinetic energy, for example in nuclear reactions. In ordinary terms, the energies are huge: Each kilogram is worth $9 \times 10^{16} J$, the amount a 1GW power station produces in 3 years!
Note that both the momentum and energy approach ∞ as u approaches c. This means that it is not possible to accelerate a particle at or beyond the speed of light. The following useful relations can be derived from the above:

$$\mathbf{u} = \mathbf{p}c^2/E$$

and

$$E^2 = p^2c^2 + m^2c^4$$

We can expand this result for large mass

$$E = mc^2 + p^2/2m - p^4/8m^3c^2 + \dots$$

giving the rest energy, usual kinetic energy, and corrections as before. We can also expand it for small mass, the “ultrarelativistic limit”.

$$E = pc + m^2c^3/2p - m^4c^5/8p^3 + \dots$$

Thus energy becomes proportional to momentum rather than its square at high energies. We could also think of a particle with zero mass, and the law

$$E = pc$$

It turns out that this can be derived from Maxwell’s (wave) theory for light, and also the quantum theory in which light has a particle nature called a photon - see later. The 4-momentum of a photon is a null (lightlike) vector, not a timelike vector.

In order to formulate conservation of energy-momentum, we must take care - what inertial reference frame are we in? For particles moving freely between collisions, there is a unique definition (diagram). In general, it is not clear how to define the conservation laws; action at a distance with Newton’s third law cannot hold.

Example: A particle A at rest splits to form two particles B and C What is the energy of particle B? We don't know and don't want to know the energy-momentum of C, so we put it on one side and square it:

$$P_A - P_B = P_C$$

$$m_A^2c^2 - 2m_A E_B + m_B^2c^2 = m_C^2c^2$$

$$E_B = \frac{m_A^2 + m_B^2 - m_C^2}{2m_A}c^2$$

Example: What is the threshold energy for the reaction $p+p \to p+p+p+\bar{p}$ for a proton striking a stationary target? If the mass of a proton is m, the final state is 4 protons at rest with respect to each other, so we have for the initial momenta,

$$P_1^2 + 2P_1 \cdot P_2 + P_2^2 = P_f^2$$

$$m^2 + 2m^2\gamma + m^2 = (4m)^2$$

$$\gamma = 7$$

The total energy is $E = 7mc^2$ and the kinetic energy is $K = 6mc^2$.

Example: A particle with kinetic energy K and mass m strikes a stationary particle of mass M. What is the kinetic energy in the centre of momentum frame? In the centre of momentum frame the total 3-momentum of the system is zero, hence the energy in this frame is simply

$$(K_{CM} + Mc^2 + mc^2)^2 = P^2c^2$$

where P is the total 4-momentum. We have

$$P^2c^2 = M^2c^4 + 2M(mc^2 + K)c^2 + m^2c^4$$

thus

$$K_{CM} = \sqrt{(M+m)^2c^4 + 2KMc^2} - (M+m)c^2$$

For small K this is

$$K_{CM} = \frac{KM}{M+m} - \frac{K^2M^2}{(M+m)^3c^2} + \ldots$$

and for large K it is

$$K_{CM} = \sqrt{2KM}c + \frac{(M+m)^2c^3}{\sqrt{8KM}} + \ldots$$

Note the slow growth with K; in practice, accelerators almost always use colliding beams, not a stationary target. The expression for K_{CM} makes sense if the moving particle is a photon with zero mass. Also, the "centre of mass" (really energy) depends on the frame; it is however unique if defined with respect to the centre of momentum frame.

6.3 Angular momentum and antisymmetry

We recall that in Newtonian mechanics, the angular momentum is defined as

$$\mathbf{l} = \mathbf{x} \times \mathbf{p}$$

In components this is

$$l_1 = x^2p^3 - x^3p^2$$
$$l_2 = x^3p^1 - x^1p^3$$
$$l_3 = x^1p^2 - x^2p^1$$

The obvious generalisation of this is

$$L^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu$$

We calculate that its derivative during free motion is zero:

$$\frac{d}{d\tau}L^{\mu\nu} = m(U^{\mu\nu} - U^{\nu\mu}) = 0$$

Also, it is obviously conserved in collisions that conserve the 4-momentum. Thus, although it takes different values in different frames, it is conserved, and is a 4-vector.
Relating this to 3-vectors, it is

$$L^{00} = L^{ii} = 0$$

$$L^{0i} = ctp^i - x^iE/c = (u^it - x^i)E/c$$
$$L^{ij} = x^ip^j - x^jp^i = e^{ijk}l_k$$

The second equation tells us, for a single particle that the particle moves with speed $\mathbf{u}$. For a collection of particles, the first term is the total momentum, and the second is the total position times energy. Dividing by the total energy, we have the RHS, where $\mathbf{u}$ is the velocity of the centre of momentum frame, and $\mathbf{x}$ is the "centre of energy", generalising the concept of centre of mass.

For the space-space part, the 3-tensor e^{ijk} equals 1 if $\{ijk\}$ is an even permutation of $\{123\}$ and -1 for an odd permutation, and zero if two or three of them are equal. It has tensor properties since it relates vectors x, p and l, but it (and $\mathbf{l}$) are actually a "pseudo-tensor" since they change sign under reflections.
We check that it satisfies

$$e^{ijk}e_{ijk} = 6$$
$$e^{ijk}e_{jkl} = 2\delta^i_l$$
$$e^{ijk}e_{klm} = \delta^i_l\delta^j_m - \delta^i_m\delta^j_l$$

The last identity leads to, for example,

$$(\mathbf{a} \times (\mathbf{b} \times \mathbf{c}))_i = e_{ijk}a_je_{klm}b_lc_m = b_i(\mathbf{a} \cdot \mathbf{c}) - c_i(\mathbf{a} \cdot \mathbf{b})$$

The second identity leads to the reverse relation

$$l_i = \frac{1}{2} e_{ijk} L^{jk}$$

Another use for the antisymmetric tensor is to compute determinants:

$$e_{ijk} A_{il} A_{jm} A_{kn} = e_{lmn} \det A$$

Applying this to a rotation matrix shows that it is the same in all orientations.

Now let us change the origin of coordinates used to define x^μ, ie the pivot. We find

$$L^{\mu\nu}(y) = L^{\mu\nu}(O) - (y^\mu P^\nu - y^\nu P^\mu)$$

thus the angular momentum depends on the pivot, as in nonrelativistic mechanics. However, in the CM frame, the total momentum is zero, so in that case, it does not. In order to find a 4-tensor which gives angular momentum independent of the pivot, we must multiply by P and antisymmetrise again!

$$S_\mu = \frac{1}{2c} \epsilon_{\mu\nu\rho\sigma} L^{\nu\rho} V^\sigma$$

here V is the 4-velocity of the CM frame, and the ϵ is now the four dimensional antisymmetric (pseudo-)tensor, $\epsilon_{0123} = 1$. It is shown to be a tensor by calculating the determinant of Lorentz transformations as above.

From an antisymmetric tensor we can now create an antisymmetric dual:

$$\tilde{F}_{\alpha\beta\gamma\delta} = \epsilon_{\alpha\beta\gamma\delta} F$$

$$\tilde{F}_{\alpha\beta\gamma} = \epsilon_{\alpha\beta\gamma\delta} F^\delta$$

$$\tilde{F}_{\alpha\beta} = \frac{1}{2} \epsilon_{\alpha\beta\gamma\delta} F^{\gamma\delta}$$

$$\tilde{F}_\alpha = \frac{1}{6} \epsilon_{\alpha\beta\gamma\delta} F^{\beta\gamma\delta}$$

$$\tilde{F} = \frac{1}{24} \epsilon_{\alpha\beta\gamma\delta} F^{\alpha\beta\gamma\delta}$$

Applying the dual operation twice gives us the original tensor back, with a minus sign because

$$\epsilon_{\alpha\beta\gamma\delta} \epsilon^{\alpha\beta\gamma\delta} = -24$$

Aside: Exterior calculus. We restrict ourselves to antisymmetric tensors with only lower indices: p-forms. The "wedge product" is an outer product with an antisymmetrisation. The "exterior derivative" is a partial derivative with an antisymmetrisation. Two exterior derivatives give zero, because two derivatives are symmetric.

If we choose our pivot to be the centre of energy as defined above, the time-space components become zero; in the CM frame this leads to

$$P_\mu L^{\mu\nu}(x_{PC}) = 0$$

the "Fokker-Synge" equation. The PC refers to "proper centroid", the centre of energy in the CM frame.

6.4 Force

In Newtonian mechanics, there are two definitions of force:

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

$$\mathbf{F} = d\mathbf{p}dt$$

The latter is known to be more useful when the mass changes, for example, in rocket problems. By analogy with the latter equation, we can then write

$$F = dP/d\tau$$

to define a 4-vector, and

$$\mathbf{f} = d\mathbf{p}dt$$

to define a 3-vector (note that this is different to NR since $\mathbf{p}$ is different). Thus

$$F = \gamma \frac{d}{dt}(E/c, \mathbf{p}) = (\gamma P/c, \gamma \mathbf{f})$$

The power P and the force transform into each other under Lorentz transformations. We construct the inner product of the 4-force and 4-velocity,

$$F \cdot U = \gamma^2 (P - \mathbf{f} \cdot \mathbf{u})$$

In the rest frame of the particle this becomes

$$F \cdot U = \frac{d}{dt}(\gamma m c^2) = c^2 \frac{dm}{dt}$$

since $d\gamma/dt = \gamma^3 \mathbf{u} \cdot \mathbf{a}/c^2 = 0$. Moving to a general reference frame, this becomes

$$F \cdot U = c^2 \frac{dm}{d\tau} = \gamma c^2 \frac{dm}{dt}$$

We define a pure force as one which does not change the (rest) mass m, ie

$$F \cdot U = 0$$

and a heatlike force as one which does not change the particle's velocity, ie

$$A = 0, \qquad F = P \frac{dm}{d\tau}$$

For example, consider a force derived from a scalar potential,

$$F_\mu = \Phi_{,\mu}$$

then

$$F \cdot U = \frac{\partial \Phi}{dx^\mu} \frac{dx^\mu}{d\tau} = \frac{d\Phi}{d\tau} \neq 0$$

in general, hence this force is not pure (unless it is zero). We find

$$mc^2 = \Phi + const$$

This type of force has been applied to the nuclear force (holding the protons and neutrons together in the nucleus).

For a pure force, we have

$$dE/dt = \mathbf{f} \cdot \mathbf{u} = \mathbf{f} \cdot d\mathbf{x}/dt$$

so the change in energy is given by the work done

$$\Delta E = \int \mathbf{f} \cdot d\mathbf{x}$$

Also we have

$$F = mA = \gamma(\mathbf{f} \cdot \mathbf{u}/c, \mathbf{f})$$

and

$$\mathbf{f} = m\gamma\mathbf{a} + \mathbf{f} \cdot \mathbf{u}\mathbf{u}/c^2$$

Note that the acceleration is only parallel to the force if the force is parallel or perpendicular to the velocity. Taking the dot product of both sides with $\mathbf{u}$, we find the parallel components as

$$f = \gamma ma + fu^2/c^2 = \gamma^3 ma$$

and the perpendicular components as

$$f = \gamma ma$$

This can be interpreted as a mass that is $\gamma^3 m$ parallel to the velocity, and γm perpendicular to the velocity.

Example: rocket moving with constant proper acceleration α, ejecting mass at speed U relative to itself. Conservation of momentum in its comoving frame gives

$$U(-dm/d\tau) = m\alpha$$

Thus

$$m_f = m_i \exp(-\alpha\tau/U)$$

We need constant acceleration expressions in terms of τ; we use the previous expression

$$\alpha = \frac{d}{dt}(\gamma v) = \frac{1}{\cosh\phi}\frac{d}{d\tau}(c\sinh\phi) = c\frac{d\phi}{d\tau}$$

Hence

$$\phi = \alpha\tau/c$$

Comparison with the previous expression gives

$$\phi = \frac{U}{c}\ln(m_i/m_f)$$

For example, a journey at v requires four acceleration phases, so even ejecting mass at the speed of light, the mass left at the end is

$$m_f = m_i \exp(-4\phi) = m_i(\frac{c+v}{c-v})^2$$

For $v = c/2$ this is $m_i/9$, for $v = 0.9c$ (to make use of some time dilation) this is $m_i/361$. Transport near c must be accomplished by outside sources!

The 4-velocity is

$$U = (\gamma c, \gamma v) = (c \cosh \alpha\tau/c, c \sinh \alpha\tau/c)$$

Note that $U^2 = c^2$. The 4-momentum is

$$P = (m_i c \exp(-\alpha\tau/U) \cosh \alpha\tau/c, m_i c \exp(-\alpha\tau/U) \sinh \alpha\tau/c)$$

For $U = c$ this becomes

$$P = ((1 + \exp(-2\alpha\tau/c))/2, (1 - \exp(-2\alpha\tau/c))/2)$$

so the 4-force is

$$F = (-\alpha \exp(-2\alpha\tau/c)/c, \alpha \exp(-2\alpha\tau/c)/c)$$

Note that here, $F^2 = 0$; compare this with $A^2 = -\alpha^2$.

6.5 Torque

We can also differentiate the angular momentum tensor to obtain the 4-torque.

$$G^{\mu\nu} = \frac{d}{d\tau} L^{\mu\nu} = x^\mu F^\nu - x^\nu F^\mu$$

The space-space part of this leads (after dividing by γ) to

$$\frac{d\mathbf{l}}{dt} = \boldsymbol{\tau} = \mathbf{x} \times \mathbf{f}$$

as before.

7 Electromagnetism

7.1 The Lorentz force

We found before that a 4-force derived from a potential could not be "pure". In order to ensure that $F \cdot U = 0$, the next most simple force is of the form

$$F_\mu = \frac{q}{c} E_{\mu\nu} U^\nu$$

with E antisymmetric. Here q is the charge, a 4-scalar that does not vary. The SI unit for charge is the Coulomb, and the charge on the electron is $1.6 \times$

$10^{-19}C$. Related units are the Ampere (Coulomb per second), and volt (Joules per Coulomb). We must compare this with the previously discovered law for motion of charges in electromagnetic fields, the Lorentz force:

$$\mathbf{f} = q(\mathbf{e} + \mathbf{u} \times \mathbf{b})$$

The SI units for electric field $\mathbf{e}$ and magnetic field $\mathbf{b}$ are Volts per metre (equals Newtons per Coulomb) and Tesla. Using $F_\mu = \gamma(\mathbf{f} \cdot \mathbf{u}/c, -\mathbf{f})$ and $U^\mu = \gamma(c, \mathbf{u})$, we get perfect agreement if

$$E_{0i} = e_i$$

$$E_{ij} = -\epsilon_{ijk} c b_k$$

This means that electric and magnetic fields transform into each other under Lorentz transformations. This is clear from two like charges at rest with respect to each other. In the frame of the charges there will be some repulsion, but in a frame moving perpendicular to the charges, this is apparently reduced due to time dilation. In nonrelativistic mechanics this smaller repulsion is explained as attraction due to magnetic fields.

We can construct the dual electromagnetic field tensor,

$$B_\mu \nu = \frac{1}{2}\epsilon_{\alpha\beta\gamma\delta} E^{\gamma\delta}$$

which has the form

$$B_{0i} = -c b_i$$

$$B_{ij} = -\epsilon_{ijk} e_k$$

There are two scalars that can be constructed from these tensors,

$$X = \frac{1}{2} E_{\mu\nu} E^{\mu\nu} = -\frac{1}{2} B_{\mu\nu} B^{\mu\nu} = c^2 b^2 - e^2$$

$$Y = \frac{1}{4} B_{\mu\nu} E^{\mu\nu} = c\mathbf{b} \cdot \mathbf{e}$$

Example: Charge in a constant electric field e in the x-direction. As the force is pure, we can integrate it to obtain a potential energy

$$V = -qex$$

If the particle starts at rest at $x = 0$, its energy is thus

$$\gamma mc^2 = mc^2 + qex$$

If it starts with a velocity in the y-direction v_i, conservation of y-momentum tells us

$$m\gamma(v_i)v_i = m\gamma v_y = m(\gamma(v_i) + qex/mc^2)v_y$$

Thus

$$v_y = \frac{\gamma(v_i)v_i}{\gamma(v_i) + qex/mc^2}$$

decreases with time, in contrast to the nonrelativistic case.

Example: Charge in a constant magnetic field b in the z-direction. As in nonrelativistic mechanics, the force is perpendicular to the momentum, so the motion is circular. The force equation becomes

$$\omega p = qub$$

The angular frequency is thus

$$\omega = qub/p$$

Note that with the nonrelativistic law $p = mu$ it is independent of the velocity, but in general is

$$\omega = qb/m\gamma$$

The radius of the circle is found using $u = \omega r$, that is,

$$r = p/qb$$

that is, it keeps increasing with the momentum of the particle; one reason why accelerators are so big.

Example: Relativistic Kepler problem. A central point charge leads to a potential energy $-k/r$ where k is a constant equal to $-qQ/4\pi\epsilon_0$, so the two constants of motion are

$$E = \gamma mc^2 - k/r$$

$$\mathbf{l} = \mathbf{r} \times \mathbf{p}$$

At the nearest and furthest points from the centre we have simply $l = rp$. Thus,

$$m\gamma c^2 = E + k/r$$

$$p^2c^2 + m^2c^4 = E^2 + 2kE/r + k^2/r^2$$

$$l^2c^2/r^2 + m^2c^4 = E^2 + 2kE/r + k^2/r^2$$

$$(m^2c^4 - E^2)r^2 - 2kEr + (l^2c^2 - k^2) = 0$$

$$r = \frac{kE \pm \sqrt{k^2E^2 - (m^2c^4 - E^2)(L^2c^2 - k^2)}}{m^2c^4 - E^2}$$

If the term $m^2c^4 - E^2 > 0$ ie the particle is bound, there will be two solutions for $L^2c^2 - k^2 > 0$. If the latter is zero, one of the solutions goes to zero, becoming negative if it is negative. Thus a charged particle with insufficient angular momentum will spiral in.

Example: transformation of fields. What is e_2 in the S' frame? We use the tensor property of $E^{\mu\nu}$:

$$e'_2 = E^{2'0'} = \Lambda_2^{2'}(\Lambda_0^{0'}E^{20} + \Lambda_1^{0'}E^{21}) = \gamma e_2 - \gamma v b_3$$

In a similar fashion we find

$$e'_1 = e_1$$

$$e'_3 = \gamma(e_3 + vb_2)$$
$$b'_1 = b_1$$
$$b'_2 = \gamma(b_2 + ve_3/c^2)$$
$$b'_3 = \gamma(b_3 - ve_2/c^2)$$

7.2 The continuity equation

We now want to derive the other half of the equations - how does the field depend on the charges? We begin by considering some charge dq in a small volume $dxdydz$. The charge density ρ is defined by

$$dq = \rho dxdydz$$

Charge is flowing into or out of the volume; the amount flowing in through the x-direction (yz planes) in time dt is

$$j^x(x)dydzdt - j^x(x+dx)dydzdt = -j^x{}_{,x}dxdydzdt$$

assuming j^x is differentiable and dx is sufficiently small. Including the other two directions we have the total amount of charge entering the volume is

$$-j^i{}_{,i}dxdydzdt$$

Since charge is conserved, this is equal to the change of charge in the volume

$$d\rho dxdydz = -j^i{}_{,i}dxdydzdt$$

thus

$$c\rho_{,0} = -j^i{}_{,i}$$

is the continuity equation, expressing conservation of charge.

How do ρ and $\mathbf{j}$ transform under the Lorentz group? It is clear that charge at rest in one frame S with density ρ_0 looks like

$$\rho' = \gamma\rho_0$$
$$\mathbf{j}' = -\gamma\mathbf{v}\rho_0$$

in S'. The γ comes from length contraction. These are just the Lorentz transformations if ρ and $\mathbf{j}$ are components of a 4-vector $J^\mu = (c\rho, \mathbf{j})$. Then the continuity equation takes on a very simple form:

$$J^\mu{}_{,\mu} = 0$$

Note that this equation is a more satisfactory expression of conservation than "Q is constant" because the latter may depend on the inertial frame.

In its rest frame, the 4-current satisfies

$$J^\mu = \rho_0 U^\mu$$

where U is the 4-velocity of the charge; thus it must hold in all inertial frames. Of course, there may be more than one current of interest, for example, positive and negative charges moving in opposite directions, leading to J^μ as a sum of the individual currents.

7.3 Inhomogeneous field equations

The electric and magnetic fields must be determined somehow from the current J^μ. We take as analogy the case of Newtonian gravity, for which the relevant equation is

$$g_{i,i} = -k\rho_m$$

where ρ_m is mass density, and k is a constant related to the Gravitational constant. So we try

$$E^{\mu\nu}{}_{,\mu} = kJ^\nu$$

We know that J must satisfy the continuity equation, so we differentiate this to get

$$E^{\mu\nu}{}_{,\mu\nu} = kJ^\nu{}_{,\nu}$$

Both sides of this equation must be zero, the LHS by antisymmetry of E, and the RHS by the continuity equation. In other words, the above equation guarantees the local conservation of charge. (In GR, the field equations guarantee the local conservation of energy and momentum).

Our field equation gives for the time component

$$\nabla \cdot \mathbf{e} = kc\rho$$

which is a Maxwell equation with $k = 1/c\epsilon_0$. ϵ_0 is the permittivity of free space, and has the value 8.85×10^{-12} Farad per metre in SI units. The space components give

$$-\frac{1}{c}\frac{\partial \mathbf{e}}{\partial t} + c\nabla \times \mathbf{b} = k\mathbf{J} = \mathbf{J}/c\epsilon_0$$

which is the other inhomogeneous Maxwell equation, if we note that

$$1/c^2\epsilon_0 = \mu_0$$

follows from $c = 1/\sqrt{\epsilon_0\mu_0}$.

7.4 Homogeneous Maxwell equations

We have only 4 equations so far, which cannot possibly determine the 6 components of field $E^{\mu\nu}$ (alternatively $\mathbf{e}$ and $\mathbf{b}$). Again, we look to Newtonian gravity,

$$g_i = -\phi_{,i}$$

and obtain the field as the gradient of a potential. In our case, we need to make sure the field is antisymmetric; this is obtained with

$$E_{\mu\nu} = \Phi_{\nu,\mu} - \Phi_{\mu,\nu}$$

and thus

$$g^{\mu\sigma}(\Phi_{\nu,\sigma\mu} - \Phi_{\sigma,\nu\mu}) = kJ_\nu$$

is 4 equations, with the 4 unknowns Φ_μ. If we write $\Phi_\mu = (\phi, -c\mathbf{w})$, the above equation reads

$$\mathbf{e} = -\nabla\phi - \frac{\partial \mathbf{w}}{\partial t}$$
$$\mathbf{b} = \nabla \times \mathbf{w}$$

In Newtonian gravity, the existence of a potential is equivalent to the equation

$$\nabla \times \mathbf{g} = 0$$

in the same way we have

$$E_{\mu\nu,\sigma} + E_{\nu\sigma,\mu} + E_{\sigma\mu,\nu} = 0$$

which translates to

$$\nabla \cdot \mathbf{b} = 0$$
$$\nabla \times \mathbf{e} + \frac{\partial \mathbf{b}}{\partial t} = 0$$

which are the homogeneous Maxwell equations. Recall the discussion about exterior calculus: we have $E = d\Phi$, $dE = dd\Phi = 0$.

If, to the potential Φ, we add the gradient of a scalar, the field is unchanged:

$$\Phi'_\mu = \Phi_\mu + \chi_{,\mu}$$

$$E'_{\mu\nu} = \Phi_{\mu,\nu} + \chi_{,\mu\nu} - \Phi_{\nu,\mu} - \chi_{,\nu\mu} = E_{\mu\nu}$$

This property is called gauge invariance; it means that only the field, not the potential is observable. It is very important in constructing quantum field theories. (We have $\Phi' = \Phi + d\chi$, $E' = d\Phi' = d\Phi + dd\chi = d\Phi = E$.) We can "fix" the gauge by imposing an additional condition, such as

$$\Phi^\mu{}_{,\mu} = 0$$

which is called the Lorentz gauge. Thus we need

$$\chi_{;\mu}{}^{;\mu} = -\Phi^\mu{}_{,\mu}$$

This linear equation has a general solution, so we can assume the Lorentz gauge condition, which deletes one term of the field equation, so

$$\Phi_{\mu}{}^{;\nu}{}_{\nu} = kJ_\mu$$

7.5 Electromagnetic waves

If we set $J_\mu = 0$ (vacuum), we have from the field equation,

$$E_{\mu\nu}{}^{;\sigma}{}_{\sigma} = 0$$

which is a wave equation with speed c, hence the speed of light. Taking a Fourier transform in space and time, we find solutions of the form

$$e^{ik_\mu x^\mu}$$

where $k_\mu = (\omega/c, -\mathbf{k})$ is a null vector ($kc = \omega$) in the direction of the wave. That is $\omega = 2\pi f$ and $k = 2\pi/\lambda$.

Substituting directly in the (3D) Maxwell equations with $J = 0$ we find that

$$\mathbf{k} \cdot \mathbf{e} = 0$$

$$\mathbf{k} \cdot \mathbf{b} = 0$$

$$\mathbf{k} \times \mathbf{e} = \omega \mathbf{b}$$

$$\mathbf{k} \times \mathbf{b} = -\omega \mathbf{e}/c^2$$

hence the orientations (polarisations), and again, $kc = \omega$.

Quantum mechanics gives the relation $P^\mu = \hbar k^\mu$ for a photon (indeed for any quantum wave/particle). $\hbar = 1.054 \times 10^{-34}$ joule seconds. Note that the velocity of the particle is $pc^2/E \leq c$, while the velocity of the wave is $\omega/k = E/p \geq c$. The energy-momentum relation

$$E^2 = p^2c^2 + m^2c^4$$

leads to the equation

$$\omega^2 = k^2c^2 + m^2c^4/\hbar^2$$

and hence

$$(\partial^\mu \partial_\mu + m^2c^2/\hbar^2)\phi = 0$$

the Klein-Gordon equation, used to model massive scalar (spin-0) particles.

Example: Doppler effect revisited. The source has 4-velocity $V = (\gamma c, \gamma v \cos\theta, \gamma v \sin\theta, 0)$ and the emitted photon has 4-momentum $P = \hbar(f/c, -f, 0, 0)$. Thus the energy of the photon in the frame of the source is

$$E' = V \cdot P = \hbar \gamma f (1 + v \cos\theta/c)$$

Thus

$$f'/f = \gamma(1 + v\cos\theta/c)$$

as before.