Spacetime Heuristics

William F. Barnes

July 16, 2015

Contents

1 Space and Time

Before the days of Albert Einstein (1879-1955), physicists weren't openly suspicious that ordinary space and might have nontrivial structure. Reasonably it was assumed that space should be trivial emptiness, and that time should pass uniformly at all points in the universe. In the pre-Einstein (a.k.a. Newtonian) view of the world, nothing could cause identical twins to age at different rates, and the trajectory of a beam of light should have nothing to do with gravity. These ideas are not only wrong, but shall be corrected by applying the most beautiful theory of physics to come out of the 20^{th} century.

1.1 Light as a Wave

The unification of electricity and magnetism through Maxwell's equations led to the idea that 'light' is a propagating pattern of electric and magnetic fields oscillating in space and time. The speed of light in vacuum is 'defined' from electrodynamics as

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \, ,$$

where ϵ and μ are called the permittivity and permeability, respectively, of a given material. (Subscript $_0$ denotes vacuum.) Measurements of c allegedly come out to $3.00 \times 10^8 \; m/s$ in vacuum.

In an attempt to classify light progapation as a type of mechanical wave, physicists speculated heavily on the required medium of travel, nicknamed the *aether*. The aether must permeate all space, and be completely stationary in the universe, with stars and galaxies moving with respect to it. It must also be extremely rigid to account for the very large value of c. Meanwhile though, it must be very transparent so as to not hinder ordinary motion. The aether was a strange and desparate hypothesis, but was completely in line with the assumptions of the day. If light is to be a wave, ther *must* be a medium - or so they thought.

Around 1887, physicists Michelson and Morley famously failed to detect any evidence of Earth's motion through the aether. (The technique of 'counting fringes' had been developed.) Instead, it was discovered that light, unlike all other types of waves, travels best in vacuum. The presence of medium is actually a hinderance to electromagnetic wave propagation. So indeed light is a wave, but no *ordinary* wave.

1.2 Einstein's Two Postulates

It's not obvious whether Einstein was aware of the Michelson-Morely experiment at the birth of special relativity. This matters little, because Maxwell's equations are the actual starting point for the theory.

Einstein was intrigued by the idea that an electric field, when viewed from a rotating reference frame, is not an electric field at all, but instead a magnetic field (and vice-versa). This was common knowledge from electrodynamics; electric fields *are* magnetic fields - the frame of reference makes all the difference.

Looking at Maxwell's equations though, there *is* no mention of reference frame. The speed of light c appears, but just who would measure this? An observer at rest? Someone on a speeding train? Maxwell's equations avail us no way to specify the reference frame. In order to have a consistent picture of the universe, Einstein was forced to tell us:

1. The laws of physics are equivalent in all inertial reference frames.

2. The speed of light is a universal constant in all reference frames.

1.3 Galilean Transformation

An *inertial reference frame* is defined as an observer (or set of observers) that is subject to no unjustified accelerations. In such a frame, events and trajectories are expressed in terms of displacement $\vec{x}$ and time t from the 'origin' $\vec{x} = 0$, $t = 0$. Consider a pair of inertial reference frames that move at a constant respective speed v. In the second frame, sometimes called the 'primed' frame, coordinates are represented with a prime symbol (a single quote), as in $(\vec{x})'$, t'.

The pre-Einstein view of the univerise is embodied in the *Galilean transformation*,

$$(\vec{x})' = \vec{x} - \vec{v}t \qquad\qquad \vec{x} = (\vec{x})' + \vec{v}t \,, \tag{1}$$

which is the way to reconcile different events as observed by reference frames in relative motion. The Galilean transformation is a rather intuitive idea - a bullet shot from a moving car gets a ground speed (not muzzle speed) bonus/penalty based on the car's motion.

1.4 Breaking the Galilean Transformation

Alice and Bob are two students standing a fixed distance apart on a straight road. If Alice shines a beam of light directly toward Bob, he shall have no trouble calculating the speed of light based on the distance and travel time of the beam. Let's say Bob's result for this experiment comes to $c = 3.00 \times 10^8\ m/s$.

Next, a car with speed v is zooming past Alice and toward Bob. Precisely as the car passes Alice, the headlights turn on, and Bob is to measure the speed of light a second time. We may suspect that Bob's new answer chould be $c + v$ because the headlights are already moving. In fact, Bob does *not* see this. The speed of light is just c again, independent of the motion of the car, and is consistent with the second postulate. Of course, the car's driver would *also* observe the headlight beams zipping away at speed c (regardless of the car's own motion), again satisfying the second postulate.

In order for all observers to agree that the speed of light is one value c, they must disagree on *something*. The Galilean transformation (1) must be *somehow* incorrect; a head-on violation of intuition.

1.5 Time Dilation and Proper Time

Consider a device made of two inward-facing parallel mirrors. A beam of light is endlessly reflecting between the mirrors, and travels along the direction normal to the faces (the light can't escape). Supposing we can know whenever the beam front encounters a face, the device at hand may be considered a type of clock, called a *light-clock*. If the distance between the mirror faces is L, then an observer holding the clock records a 'tick' every c/L seconds. Defining the time between successive ticks as Δt, we write

$$L = c\Delta t \,.$$

Imagine that a light-clock exists in a reference frame that moves at speed v with respect to a stationary observer. The clock's motion is perpendicular to the beam's direction of travel. According to the observer, the beam front must move through a distance of

$$\sqrt{L^2 + (v\Delta t')^2} = c\Delta t' \,,$$

where $\Delta t'$ is the observed interval between successive ticks. In the above, note that speed c does not need a prime ($'$) superscript by Einstein's second postulate.

Eliminate L between the two above equations to arrive at a profound relation

$$\Delta t' = \frac{1}{\sqrt{1 - v^2/c^2}} \Delta t \,, \tag{2}$$

the famous *time-dilation* effect. Equation (2) shows that reference frames in relative motion cannot have the same sense of time. Notice that the specific construction of the light-clock is irrelevant. Only the speed v between the moving observer and the stationary clock is important.

The combination $(1 - v^2/c^2)^{-1/2}$ occurs often in special relativity and is denoted by the greek letter γ, pronounced 'gamma', which is always greater than 1. The time Δt measured in the clock's own rest frame is defined as the *proper time*, denoted by the greek letter τ ('tau'). It follows from equation (2) that the proper time τ is the shortest of all times $\Delta t'$ describing the same interval. In summary we write

$$\Delta t' = \gamma \tau \qquad\qquad dt' = \gamma d\tau \,. \tag{3}$$

Problem 1: The mean lifetime τ of a subatomic muon (denoted μ) is known to be 2.2 microseconds. Suppose an atmospheric muon is generated at altitude $6000m$, and is detected at sea level just as it decays. Assuming the muon's motion was straight down, what is the speed of the muon as a fraction of c? Answer: $0.994c$

1.6 Length Contraction

It should be no surprise that the concept of 'length' is in need of revision to accomodate Einstein's postulates. Consider a fast-flying insect moving at speed v in one dimension (flying parallel to a ruler if you like). A stationary observer watches the insect traverse a distance L_p due to its speed v over the time interval $\Delta t'$. According to the insect, the perceived distance traveled L' occurs over the proper time interval τ. We thus have two equations

$$L_p = v\Delta t' \qquad\qquad L' = v\tau,$$

which can be simplified to yield the famous *length-contraction* effect:

$$L' = \frac{L_p}{\gamma} \tag{4}$$

We now define the rest-frame length of an object to be the *proper length*, denoted L_p. Equation (4) tells us that a moving observer sees the world flatten along the direction of motion. This is also true in reverse - objects in motion appear flatten in the direction of their motion. Note this effect is not due to some kind of pressure or extra force. That is, nobody feels length contraction - space is the thing that is flexing.

Problem 2: A muon in the atmosphere is moving toward the ground at speed $0.994c$. When the muon has altitude $6000m$ as observed from the ground, what is the apparent altitude in the muon's rest frame? Answer: $656m$

1.7 Lorentz Transformation

The Galilean transformation clearly does not account for time dilation or length contraction. The 'object' that does satisfy Einstein's postulates is called the *Lorentz transformation*, which is essentially a nonlinear rescaling of (1). For one-dimensional motion, the Lorentz transformation reads

$$x' = \gamma\,(x - vt) \qquad\qquad x = \gamma\left(x' + vt'\right)\ , \tag{5}$$

where solving for the pair t, t' delivers

$$t' = \gamma\left(t - vx/c^2\right) \qquad\qquad t = \gamma\left(t' + vx'/c^2\right)\ . \tag{6}$$

Variables with a prime ($'$) superscript correspond to measurements made by a moving observer. The unprimed reference frame is considered stationary. If the relative velocity is constant, the reference frames are called *inertial*. General treatment of an accelerated reference frame is beyond the scope of special relativity.

Equations (5) and (6) apply to one-dimensional, constant-velocity motion only, leaving the y- and z- spatial dimensions unchanged. Since the quantity v/c occurs frequently, define the dimensionless parameter

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

which may be greater than 0 but never greater than 1. In matrix form, the Lorentz transformation for motion along x, also known as a *boost* along x, reads

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}\ . \tag{7}$$

Note that a boost along the x-direction noes not warp the space in the y- and z-directions.

Problem 3: Write the Lorentz transformation matrix for a boost along the y direction. Repeat for a boost along z.

Problem 4: Show that the Lorentz transformation for constant-velocity motion in an arbitrary direction is given by

$$\begin{bmatrix} ct' \\ x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1 + (\gamma-1)\,\beta_x^2/\beta^2 & (\gamma-1)\,\beta_x\beta_y/\beta^2 & (\gamma-1)\,\beta_x\beta_z/\beta^2 \\ -\gamma\beta_y & (\gamma-1)\,\beta_y\beta_x/\beta^2 & 1 + (\gamma-1)\,\beta_y^2/\beta^2 & (\gamma-1)\,\beta_y\beta_z/\beta^2 \\ -\gamma\beta_z & (\gamma-1)\,\beta_z\beta_x/\beta^2 & (\gamma-1)\,\beta_z\beta_y/\beta^2 & 1 + (\gamma-1)\,\beta_z^2/\beta^2 \end{bmatrix} \begin{bmatrix} ct \\ x \\ y \\ z \end{bmatrix}\ .$$

Answer: Let $\vec{r} = x\hat{\imath} + y\hat{\jmath} + z\hat{k}$ so that $\vec{v} = (dx/dt)\hat{\imath} + (dy/dt)\hat{\jmath} + (dz/dt)\hat{k}$. Recast the $\vec{r}$-vector into its 'parallel-to-v' and 'perpendicular-to-v' components

$$\vec{r} = \vec{r}_{\parallel v} + \vec{r}_{\perp v}\ .$$

Note next that only the $r_{\parallel v}$-component of the $\vec{r}$-vector is warped by the Lorentz transformation. Equations (5) and (6) generalize to

$$\vec{r}\,' = \vec{r}_{\perp v} + \gamma\left(\vec{r}_{\parallel v} - \vec{v}t\right) \qquad\qquad t' = \gamma\left(t - \vec{r}\cdot\vec{v}/c^2\right)\ .$$

The $\vec{r}'$ equations needs to be expressed purely in terms of $\vec{r}$. Replace all instances of $\vec{r}_{\perp v}$ with $\vec{r} - \vec{r}_{\parallel v}$. Note next that $\vec{r}_{\parallel v}$ is parallel to $\vec{v}$, making $\vec{r}_{\parallel v} = \vec{r} \cdot \vec{v}/v$. Thus:

$$\vec{r}' = \vec{r} + \left(\frac{\gamma - 1}{v^2}\vec{r}\cdot\vec{v} - \gamma t\right)\vec{v}$$

Note that $\vec{\beta} = c^{-1}(v_x\hat{\imath} + v_y\hat{\jmath} + v_z\hat{k})$ and the column vector $\vec{\beta}^T$ is the transpose of $\vec{\beta}$. In block matrix form, the t' and $\vec{r}'$ equations may be written as

$$\begin{bmatrix} ct' \\ \vec{r}' \end{bmatrix} = \begin{bmatrix} \gamma & -\gamma\vec{\beta}^T \\ -\gamma\vec{\beta} & I + (\gamma - 1)\,\vec{\beta}\vec{\beta}^T/\beta^2 \end{bmatrix} \begin{bmatrix} ct \\ \vec{r} \end{bmatrix} ,$$

where I is the identiy matrix. This result is equivalent to the relation we set to prove.

1.8 Relativity of Simultaneous Events

An *event* is a precise location in space and time according to some observer. An *interval* is the separation between two events in space-time. Two events separated by distance Δx and time Δt can be described in a Lorentz-boosted frame by taking the 'differential' version of (5) and (6). Doing so we write:

$$\Delta x' = \gamma\left(\Delta x - v\Delta t\right) \qquad \Delta x = \gamma\left(\Delta x' + v\Delta t'\right) , \tag{8}$$

$$\Delta t' = \gamma\left(\Delta t - v\Delta x/c^2\right) \qquad \Delta t = \gamma\left(\Delta t' + v\Delta x'/c^2\right) \tag{9}$$

It is now apparent that the idea of 'simultaneous' events is now dependent on reference frame. Consider two events that, as observed in a stationary reference frame, occur at the same instant in time, but are separated by some distance apart. (This may be the flashing of lights on airplane wings, with the pilot being the 'stationary' observer.) In such a reference frame, we may describe the interval between the two events by $\Delta t = \tau = 0$ and $\Delta x \neq 0$. In some different inertial reference frame, the two events (flashes) are *not* observed to be simultaneous. Solving for $\Delta t'$ in (8) or (9) tells us

$$\Delta t' = -\gamma v \Delta x/c^2 ,$$

which is impossible to attain with the Galilean transformation (1).

Consider a new set of events that, as observed in a stationary reference frame, occur at two different points in time ($\Delta t = \tau \neq 0$), yet at the same location ($\Delta x = 0$). It follows that

$$\Delta x' = -\gamma v \Delta t ,$$

with the presence of the γ factor crystalizing the fact that time and space are mixed in a nontrivial fashion.

1.9 Addition of Velocities

Suppose there exist two reference frames S and S' moving with some relative speed v in the x-direction. Suppose also that a fast moving particle, also moving along the x-direction, is observed to move at speeds u_x and u_x', respectively. For definiteness, assume that $u_x > v$ as observed in frame S.

The 'usual' assumption is to take $u_x = v + u'_x$, but the differential Lorentz transformation, equations (8) and (9), contradict this. Computing $u_x = \Delta x / \Delta t$ we obtain:

$$u_x = \frac{v + u'_x}{1 + v u'_x / c^2} \qquad\qquad u'_x = \frac{u_x - v}{1 - v u_x / c^2} \tag{10}$$

The story goes on - if the particle's motion were instead along the y-direction, the velocities are still affected, despite the boost being along x. In such a case we have $\Delta y = \Delta y'$, and relating u_y to u'_y results in:

$$u_y = \frac{u'_y}{\gamma \left(1 + v u'_x / c^2\right)} \qquad\qquad u'_y = \frac{u_y}{\gamma \left(1 - v u_x / c^2\right)} \tag{11}$$

1.10 Rapidity

Since velocities are not directly additive, it is natrual to wonder just which quantity *is* additive under two Lorentz boosts. To answer this we work backward from the anticipated result

$$\Phi = \phi + \phi' \ ,$$

where each parameter ϕ is called the *rapidity*, and relates to each speed by some unknown function as follows:

$$u_x = f\left(\phi + \phi'\right) \qquad\qquad v = f\left(\phi\right) \qquad\qquad u'_x = f\left(\phi'\right)$$

Next, take the 'addition of velocities' formula (10) and replace all instances of speed (u_x, v, u'_x) with the corresponding f-representations:

$$f\left(\phi + \phi'\right) = \frac{f\left(\phi\right) + f\left(\phi'\right)}{1 + f\left(\phi\right) f\left(\phi'\right)}$$

In this form, it is highly evident that the only function suitable for the job of $f\left(\right)$ is the hyperbolic tangent $\tanh\left(\right)$, because

$$\tanh\left(\phi + \phi'\right) = \frac{\tanh\left(\phi\right) + \tanh\left(\phi'\right)}{1 + \tanh\left(\phi\right) \tanh\left(\phi'\right)} \ ,$$

and no other function obeys such an identity. Conclude that speed is related to rapidity by:

$$u_x = c \tanh\left(\Phi\right) \qquad\qquad v = c \tanh\left(\phi\right) \qquad\qquad u'_x = c \tanh\left(\phi'\right) \tag{12}$$

<u>Problem 5</u>: Show that the simplest formula that relates the rapidity ϕ to the gamma factor γ is given by

$$\cosh\left(\phi\right) = \gamma \ . \tag{13}$$

Also show that

$$\sinh\left(\phi\right) = \gamma v / c \ . \tag{14}$$

<u>Problem 6</u>: Show that the effective Lorentz transformation matrix for two boosts in the x-direction is given by:

$$\begin{bmatrix} \cosh\left(\phi + \phi'\right) & -\sinh\left(\phi + \phi'\right) & 0 & 0 \\ -\sinh\left(\phi + \phi'\right) & \cosh\left(\phi + \phi'\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{15}$$

1.11 Spacetime Interval and Causality

Observers in different inertial reference frames don't generally agree on the coordinates or duration of events. It's natural to wonder if such observers agree on *anything* at all. As it turns out, there do exist quantities, called *invariants*, that are the same in all reference frames. One invariant is called the *spacetime interval*, denoted ΔS^2, and is written as

$$\Delta S^2 = -\left(c\Delta t\right)^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 \tag{16}$$

for any pair of events. The same pair of events, as observed in a coving reference frame, satisfies

$$\left(\Delta S'\right)^2 = -\left(c\Delta t'\right)^2 + \Delta x'^2 + \Delta y'^2 + \Delta z'^2 \,.$$

The spacetime interval is called invariant because

$$\Delta S^2 = \left(\Delta S'\right)^2 \,. \tag{17}$$

<u>Problem 7</u>: Prove equation (17) using equations (8) and (9).

The spacetime interval is actually equivalent to $-c^2$ multiplied by the square of the proper time $\Delta\tau$. To see this, start with equation (16) and replace the quantity $d\vec{x}$ with $\vec{v}\Delta t$, giving

$$\Delta S^2 = \frac{-c^2\Delta t^2}{\gamma^2}.$$

Replace γ^2 with the differential version of (3), and arrive at

$$\Delta S^2 = -c^2\Delta\tau^2 \,. \tag{18}$$

Corollary

The spacetime interval between any pair of events is equal to $-c^2$ times the square of the rest-frame time between those events. I remind you that the rest-frame time, a.k.a. proper time, is smaller than all other $\Delta t'$ corresponding to the same pair of events.

<u>Problem 8</u>: The trajectory of particle moving at speed $c/2$ in the x-direction may be parameterized by ($t = 2\lambda/c$, $x = \lambda$). Using equation (18), show (i) that λ is not the proper time, and (ii) that the proper time is equal to $\tau = \sqrt{3}\lambda$.

Causality

Despite the fact that observersin relative motion generally disagree on measurements of the same events, it should be obvious that the notion of 'cause-and-effect' cannot be violated. For example, if event A causes event B, there is no reference frame where one could observe B causing A.

We classify the causal relationship between any pair of events as *spacelike*, *timelike*, or *lightlike* by the sign of ΔS^2 as follows:

- Spacelike events have $\Delta S^2 > 0$. In this case, the time span between the events is very short, and their spatial separation is large. It follows that these events are not causally connected - the first event could not influence the second event.

- Timelike events have $\Delta S^2 < 0$. In this case, the events are sufficiently close togaether in space, and the time Δt is sufficiently large that the first event could have influenced the second event.

- Lightlike events have $\Delta S^2 = 0$. A pair of lightlike events are situated such that a ray of light emitted from event A arrives precisely at event B.

1.12 Twin Paradox

Problem 9: Consier a 'twin paradox' in which one twin, named Albert, stays at home on Earth. The seond, named Barney, travels 5 light-years straight out into space at speed $c/2$ and returns back to Earth at speed $c/4$. Determine (i) how much Albert has aged during Barney's trip, and (ii) how much Barney has aged during his own trip. Answer:

$$\Delta T_A = 30\,yr \qquad\qquad \Delta T_B = 5\sqrt{3}\,yr + 5\sqrt{15}\,yr \approx 28\,yr$$

Problem 10: To answer part (ii) in the question above, did you use Lorentz transformation equations, or did you exploit invariance in the spacetime interval? Whichever you chose, attain the same answer by the other method (even if you know the methods to be equivalent).

Problem 11: Barney gets a new, more powerful spaceship. With what speed would he have to make the same round trip in order to have aged only 1% as Albert? Answer:

$$v = c\sqrt{1-(1\%)^2} \approx c \times 0.99994999875...$$

2 Relativistic Mechanics

2.1 Force and Momentum

Consider a uniformly-accelerated particle that is observed in an intertial reference frame S. According to observers in frame S, the speed of the particle at any given time t is $u_x(t)$. Suppose another observer O' in a covmoving inertial frame follows the particle for a given instant and observes the speed to be $u'_x(t')$. Suppose further that the particle does *not* move relativistically according to O'. (That is, the gamma factor between O' and the particle is indistinguishable from unity.) According to O', the force on the partcle is

$$F' = m\frac{du'_x}{dt'} = m\frac{du'_x}{dt}\frac{dt}{dt'} = m\frac{d}{dt}\left(\frac{u_x - v}{1 - vu_x/c^2}\right)\frac{d}{dt'}\left(\gamma t' + \gamma v x'/c^2\right),$$

where the chain rule, along with equations (6) and (10), has been used.

After evaluating the derivatives above, make the crucial step: let the frame O' *equal* (catch up to) the frame of the particle. That is, we let $v = u_x$, $u'_x = 0$, and $F' = F$. Therefore,

$$F = \frac{m}{(1-u_x^2/c^2)^{3/2}}\frac{du_x}{dt} = m\frac{d}{dt}\left(\gamma u_x\right)\,,$$

where the relation

$$\frac{1}{(1-z^2)^{3/2}} = \frac{d}{dz}\frac{z}{\sqrt{1-z^2}}$$

has been used.

Apparently, Newton's second law retains its simplicity after accounting for special relativity. In effect, the only change is $u \to \gamma u$. Generalizing to three dimensions, the *relativistic force* is written as

$$\vec{F} = m\frac{d}{dt}(\gamma\vec{u}) \ , \tag{19}$$

and we immediately pick out the *relativistic momentum* to be

$$\vec{p} = m\gamma\vec{u} \, . \tag{20}$$

Problem 12: For small speeds u, use binomial expansion to show that the formula $p = m\gamma u$ reduces to the Newtonian formula for momentum, and find the next-order corrections. Answer: $p = mu\left(1 + \beta^2/2 + 3\beta^4/8 + ...\right)$

Problem 13: A particle of mass m is at rest at time $t = 0$ and is subject to a uniform force F for times $t > 0$. Determine the equations of motion. Answer:

$$v(t) = \frac{c}{\sqrt{1 + m^2c^2/F^2t^2}} \qquad\qquad x(t) = \frac{c}{F}\sqrt{F^2t^2 + m^2c^2} - \frac{mc^2}{F} \tag{21}$$

2.2 Kinetic Energy and Rest Energy

The kinetic energy of a relativistically-moving particle is not $\frac{1}{2}mv^2$. (In fact, this definition isn't true in quantum mechanics either, so you may as well dispense with the formula.) The true kinetic energy T is determined by calculating the work required to accelerate a particle to relativistic speed from rest (assome no potential energy). That is,

$$T = \int dW = \int \vec{F}\cdot d\vec{r} = \int m\frac{d}{dt}(\gamma\vec{u})\cdot\vec{u}dt = m\int u\frac{d}{dt}(\gamma u) \ , \tag{22}$$

where $\vec{F}$ has been replaced by (19), $\vec{u}\cdot d\vec{u}/dt = u(du/dt)$ has been used. After the dust settles, we find that the kinetic energy of a relativistic particle is:

$$T = mc^2(\gamma - 1) \tag{23}$$

Problem 14: Prove that $\vec{u}\cdot d\vec{u}/dt = u(du/dt)$, and then use integration by parts to derive (23) starting with (22).

Problem 15: For small speeds v, use binomial expansion to show that equation (23) reduces to the Newtonian formula $(1/2)mv^2$, and find the next-order corrections. Answer: $T = (1/2)mv^2 + (3/8)mv^4/c^2 + (5/16)mv^6/c^4 + ...$

Equation (23) is a loaded formula. First notice that the speed dependence is contained in γ, yet one of the terms in T has a γ-independent term, namely mc^2. Einstein noticed this strange fact, and reasoned that mc^2 is an energy amount that arises simply due to the existence of the mass m. The lowest energy state (no velocity) of a particle is

$$E_0 = mc^2 \ ,$$

which is undoubtedly the most famous equation in the world. We interpret E_0 as the *rest energy* of the particle. The total energy of the particle is written

$$E = T + E_0 = \gamma mc^2 \ . \tag{24}$$

Perhaps the most useful formula to come out of special relativity is the *momentum-energy* relation. Square both sides of $\vec{p} = m\gamma\vec{u}$ and replace u^2 with $1 - 1/\gamma^2$ to arrive at

$$E^2 = p^2c^2 + m^2c^4 \ . \tag{25}$$

Problem 16: Starting with euqation (25), prove:

$$\frac{dE}{dt} = \frac{d\vec{p}}{dt} \cdot \vec{v} \tag{26}$$

2.3 Collision of Identical Masses

When two separate objects with masses m_1, m_2 collide and fuse togaether to form a heavier object, the new mass M is *not* the sum $m_1 + m_2$. To illustrate, suppose two identical particles of mass m and equal speeds collide head-on, and this reaction is observed in the center-of-mass or *COM* reference frame. The total energy of the system is

$$E = 2 \times \left(T + mc^2\right) \ ,$$

where T is the kinetic energy of each particle. Suppose that after the collision, a single particle of mass M remains with zero speed, making $E = Mc^2$. The mass of the resultant particle is therefore

$$M = 2m + 2T/c^2 \ .$$

2.4 Decay of a Massive Particle

The problem also applies in reverse - suppose a stationary particle of mass M spontaneously decays into two non-identical particles of mass m_i, energy E_i, and momentum p_i with $i = 1, 2$. Conservation of energy and momentum allow us to write

$$Mc^2 = \sqrt{p_1^2c^2 + m_1^2c^4} + \sqrt{p_2^2c^2 + m_2^2c^4} \qquad\qquad \vec{p}_1 + \vec{p}_2 = 0 \ ,$$

which, after a bit of algebra condenses into

$$M^2 = m_1^2 + m_2^2 + 2\left(\frac{E_1E_2}{c^4} - \frac{\vec{p}_1 \cdot \vec{p}_2}{c^2}\right) \ .$$

The energies E_1 and E_2 can also be isolated, resulting in:

$$E_1 = c^2\frac{M^2 + m_1^2 - m_2^2}{2M} \qquad\qquad E_2 = c^2\frac{M^2 + m_2^2 - m_1^2}{2M} \tag{27}$$

Problem 17: Prove equation(s) (27).

2.5 Massless Particles, Photons

The energy of a relativistically-moving particle is $E = \sqrt{p^2c^2 + m^2c^4}$, which was built on the assumption that the mass m is greater than zero. Let's now conceive a new kind of particle that has $m = 0$, but we apply the energy formula *anyway*. It follows that massless particles obey $E = pc$. Meanwhile we know

$$E = \gamma mc^2 \qquad\qquad p = m\gamma v \ ,$$

and immediately find that $v = c$ for massless particles. That is, *massless particles travel at the speed of light.*

Special relativity has deliviered evidence of something curious - namely because massless particles are non-Newtonian objects. The question is, what *is* a massless particle? We assign to them a name called *photons*, which are interpreted as localized packets of electromagnetic radiation. Of course, electromagnetic radiation is supposed to be wave-like, not particle-like. It turns out that both interpretations are correct. This is the famous *duality of light.*

Around the early 1920's, experiments were showing that the energy of a photon (light) is directly proportional to the frequency (color), up to a proportionality factor h called *Planck's constant.* We gain the equation

$$E = hf \,, \tag{28}$$

where f is the frequency in Hz and $h = 6.626 \times 10^{-34} Js$. Define the new constant $\hbar$ (pronounced 'h-bar') as $\hbar = h/2\pi$. Noting that frequency and wavelength λ relate by $c = f\lambda$, and further that the wavenumber k is defined as $k = 2\pi/\lambda$, the momentum of a photon must be

$$p = \hbar k \,. \tag{29}$$

2.6 Momentum-Energy Lorentz Transformation

The 'ordinary' Lorentz transformation (5), (6) is cast in terms of the variables $t, \vec{x}, t', \vec{x}'$. Another convenient set of variables is $E, \vec{p}, E', \vec{p}'$, as measured in two reference frames S and S' that have relative speed v (assume one dimensional motion for simplicity). To begin, write the energy E' of a particle as

$$E' = \gamma' mc^2 = mc^2 \Big/ \sqrt{1 - \frac{1}{c^2}\left(\frac{u_x - v}{1 - vu_x/c^2}\right)^2} \,,$$

where u_x' has been replaced by the 'addition of velocities' formula (10). Noting that

$$E = \frac{mc^2}{\sqrt{1 - u_x^2/c^2}} \qquad\qquad p = \frac{mu_x}{\sqrt{1 - u_x^2/c^2}} \,,$$

the above simplifies to

$$E' = \gamma \left(E - pv\right) \,. \tag{30}$$

Repeating the same exercise for the momentum, start with $p' = \gamma' m u_x'$, and simplify to attain

$$p' = \gamma \left(p - Ev/c^2\right) \,. \tag{31}$$

Problem 18: Prove equations (30) and (31). Also attain the inversion of the momentum-energy Lorentz transformation. That is, solve for E and p in terms of E' and p'.

Problem 19: Use both forms of the Lorentz transformation to prove that the quantity $px - Et$ is invariant. In other words, show that

$$p'x' - E't' = px - Et \tag{32}$$

for one-dimensional motion.

2.7 Relativistic Doppler Shift

A fudamental postulate of special relativity is that the speed of a photon is an invariant. This does not mean that the frequency (or energy) is also invariant. Indeed, the reference frame determines the frequency of an observed photon. To derive the *relativistic doppler shift*, start with the invariant (32), and replace all instances of E and p with (28) and (29), respectively. This delivers a relationship between observed frequencies f and f':

$$f = f' \sqrt{\frac{1 + v/c}{1 - v/c}} \, , \tag{33}$$

where v is the relative refernce frame velocity. To interpret the minus-sign placement in (33), remember that a photon moving toward the observer has increased energy (blueshift), and a photon moving away has diminished energy (redshift).

3 Minkowski Space and Four-Vectors

3.1 Spacetime Diagram

As usual in physics, plots and graphs are a great visual aid for keeping track of motion and many other things. In special relativity, we introduce a new idea called the *spacetime* a.k.a. *Minkowski diagram.* The Minkowski diagram is a plot with space (say x) on the horizontal axis, and the vertical axis has units ct, as depicted in Figure 1.

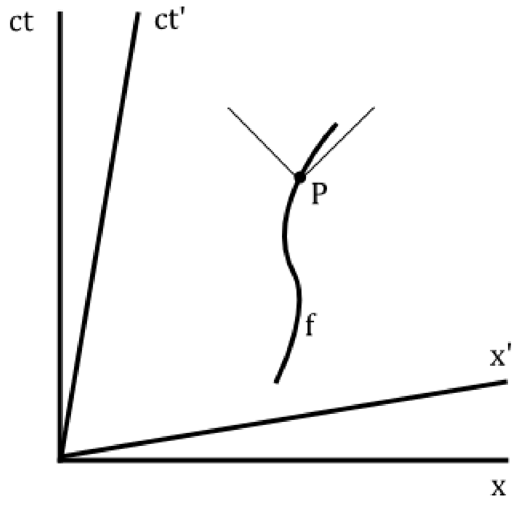

Figure 1: Minkowski diagram in 1+1 dimensions with light cone.

The trajectory f of some particle is a continuous line, called a *worldline* in the spacetime diagram. The velocity of the particle at a given instant corresponds to the slope of the worldline. The axes are scaled such that a 45-degree slope corresponds to a speed of c. It follows that no massive particle can have a worldline whose slope exceeds 45°.

A primed observer may observe the same particle in terms of coordinates x', t', but we need not draw a second worldline. Instead, we draw, on the same diagram, a new set of axes x', ct'. Each primed axis is tilted off of the x- and t-axes, respectively, by one and the same angle that has to do with the speed v between the two reference frames. Note that the reference frames are overlapping at $x = x' = t = t' = 0$.

At point P suppose that the particle emits a flash of light that propagates along $\pm x$. Emerging from P is a pair of 45° lines that trace the beam front - this is called the *light cone*. Notice that event P has no possible way of influencing future events that occur outside of its light cone - such events are 'spacelike' with P. On the other hand, future events inside the light cone are 'timelike' with P. Future events that live on the light cone are 'lightlike' with P.

Problem 20: Prove that the angle between the ct and ct' axes in a spacetime diagram is equal to $\tan^{-1}(v/c)$. Hint: Lorentz-transform the point $(ct' = 1, x' = 0)$ into unprimed coordinates.

3.2 Position Four-Vector

Special relativity deals with inertial reference frames that are far away from any source of gravitation - locations called *Minkowski space*. Excluding the time dimension, we simply say *flat space*. The physical universe is said to be $1+3D$, meaning there is one time and three spatial dimensions.

It makes the most sense to treat time as a spatial variable, which is actually what has already happened in the Lorentz transformation. To formalize the idea, introduce the *position four-vector*

$$x^\mu = (ct, x, y, z) \ , \tag{34}$$

which specifies an event at a particular location in spacetime. The index μ equals 0, 1, 2, and 3 such that $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$. Not just any combination of variables can constitute a four-vector; it must transform by (some version of) the Lorentz transformations.

Four-vectors always exist in some basis. That is, the expression A^μ doesn't refer to the entire vector A, but only to the μ^{th} component of A in the basis $\vec{e}_{(\mu)}$. Any vector (three-, four- or otherwise) is a linear combination of amplitudes and basis vectors according to the scheme

$$A = \sum_{\mu=1}^{4} A^\mu \vec{e}_{(\mu)} \ , \tag{35}$$

where in general the number 4 is replaced with the *dimension* of the space in which A lives.

Problem 21: Consider a worldline $x^\mu(\lambda)$ that goes across the event $E = (0, 0, 1m, 0)$, parameterized as

$$x^0(\lambda) = \lambda \qquad\qquad x^1(\lambda) = \lambda/2$$

$$x^2(\lambda) = (1m)\sqrt{1 - (\lambda/1m)^2} \qquad\qquad x^3(\lambda) = 0 \ .$$

(i) What value of λ corresponds to event E? (ii) Compute the components of the *tangent vector* $dx^\mu/d\lambda$ to the worldline $x^\mu(\lambda)$ at event E. (iii) Compute the velocity dx^j/dt of a particle that goes along the worldline $x^\mu(\lambda)$ at event E. Answer:

$$\lambda = 0 \qquad\qquad \left.\frac{dx^\mu}{d\lambda}\right|_E = (1m, 0.5m, 0, 0)$$

$$\left.\frac{dx^j}{dt}\right|_E = \left(\frac{c}{2}, 0, 0\right)$$

Problem 22: Suppose the worldline in the previous problem is expressed in a frame that is boosted with speed v along the x^1 direction in accordance with (7). Compute the nonzero components of the tangent vector $(dx^\mu)'/d\lambda$ in the primed reference frame. Answer:

$$\frac{\left(dx^0\right)'}{d\lambda} = \frac{d}{d\lambda}\left(\frac{\lambda\left(1-v/2c\right)}{\sqrt{1-v^2/c^2}}\right) = \gamma\left(1-\frac{\beta}{2}\right) \qquad \frac{\left(dx^1\right)'}{d\lambda} = \gamma\left(\frac{1}{2}-\beta\right)$$

3.3 Velocity Four-Vector

The position four-vector is defined such that the square of dx^μ is invariant. Taking equation (16) as the structure of the 'square' of dx^μ, we notice

$$||dx \cdot dx|| = dS^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \tag{36}$$

is the same in all inertial reference frames. Note the minus sign on the dt^2 term. This is a non-obvious generalization of the dot product that we'll call the 'Minkowski dot product', also known as the 'Lorentz-invariant inner prouct'.

The derivative of the position four-vector with respect to the proper time yields the *four-velocity*:

$$U^\mu = \left(c\frac{dt}{d\tau}, \frac{d\vec{x}}{d\tau}\right) = (\gamma c, \gamma\vec{v}) \tag{37}$$

The identity $dt/d\tau = \gamma$ has been derived from (3). Computing the square of U by the Minkowski dot product, we see

$$||U \cdot U|| = -\gamma^2c^2 + \gamma^2v^2 = -c^2 ,$$

which is surely invariant.

The *four-momentum* is defined by multiplying the four-velocty by the rest mass m, condensing the energy and the momentum into one object P^μ as

$$P^\mu = mU^\mu = \left(\frac{E}{c}, \vec{p}\right) , \tag{38}$$

which obeys the momentum-energy Lorentz transformation. Noting that $||P \cdot P||$ is automatically equal to the invariant $-m^2c^2$, applying the Minkowski dot product to (38) recovers the energy-momentum formula, namely $E^2 = p^2c^2 + m^2c^4$, originally written as equation (25).

Problem 23: Using the Minkowski dot product, evaluate $||x \cdot P||$ to recover the invariant (32). In other words, show that

$$||x \cdot P|| = px - Et .$$

Problem 24: Consider a photon moving at speed c in one dimension. Show, using the Minkowski dot product, that the square of the photon's four-momentum is identically zero.

3.4 Relativistic Collisions

From the standpoint of Newtonian kinematics, the microscopic world is a hectic game of billiards, but nothing truly exotic ever happens. In any system, the amount of matter should be a rigid constant which particles are free to exchange energy and momentum. Accounting for special relativity however, we'll see that the amount of matter in any system is *not* constant. Particle creation and annihilation is routine in the subatomic world - for instance, you can actually make matter vanish in exchange for a great amount of energy.

Photo-π^0 Production

Consider a single 'target' proton of mass m_p that is stationary with respect to a laboratory reference frame. An incoming photon of energy E_γ collides with the proton and is absorbed. A π^0 particle is created in the collision, which carries away momentum $\vec{k}_\pi$. The proton also recoils with momentum $\vec{k}_p$.

In four-vector notation, this collision, known as 'photo-π^0 production', may be written as

$$\gamma + M = \pi + P \,, \tag{39}$$

where the individual four-vectors are

$$\gamma = \left(\frac{E_\gamma}{c}, \vec{k}_\gamma\right) \qquad M = (m_p c, 0) \qquad \pi = \left(\frac{E_\pi}{c}, \vec{k}_\pi\right) \qquad P = \left(\frac{E_p}{c}, \vec{k}_p\right) \,.$$

In accordance with Einstein's first postulate, photo-π^0 production must obey conservation of energy and momentum. If four-vectors are useful at all, we should explicitly recover energy and momentum conservation statements from $\gamma + M = \pi + P$.

The energy conservation statement comes from multiplying (Minkowski-dotting) both sides of (39) by the four-vector M, resulting in

$$E_\gamma + m_p c^2 = E_\pi + E_p \,,$$

as expected. A similar trick comes from Minkowski-dotting the four-vector γ into both sides of (39). If we denote the x-axis as the beamline of the photon, then the π^0 particle and recoil proton fly off at angles ϕ and θ, respectively. It follows that

$$k_\gamma = k_\pi \cos\phi + k_p \cos\theta \,,$$

correctly stating the conservation of momentum in the x-direction.

Problem 25: Starting from equation (39), derive the full statement of momentum conservation, $\vec{k}_\gamma = \vec{k}_\pi + \vec{k}_p$. Hint: Square both sides of $\gamma + M - P = \pi$, and then square both sides of $\gamma + M - \pi = P$, and then add the results. Use energy conservation along the way.

Compton Scattering

Problem 26: Consider a single 'target' proton of mass m that is stationary with respect to a laboratory reference frame. An incoming photon of energy E_i collides with the proton. The photon scatters with energy E_f at an angle θ relative to the beam axis, and the proton carries some energy away to make $E_i > E_f$. In four-vector notation, you may write

$$\gamma_i + M = \gamma_f + P \ .$$

Derive the Compton scattering formula:

$$\frac{1}{E_f} - \frac{1}{E_i} = \frac{1}{mc^2}\left(1 - \cos\theta\right) \tag{40}$$

Problem 27: Suppose the recoil proton in a Compton scattering event emerges with kinetic energy T and at angle ϕ with respect to the beam axis. Derive the relation:

$$\cos\phi = \frac{T\left(1 + mc^2/E_i\right)}{\sqrt{\left(mc^2 + T\right)^2 - m^2c^4}}$$

The *Compton wavelength* is a convenient way to state the effective 'size' of the target proton. Using the Planck equation (28) to replace the photon energies, namely $E = hf$ with $f = c/\lambda$ being the frequency and λ being the wavelength, the Compton scattering formula becomes

$$\lambda_f - \lambda_i = \frac{h}{mc}\left(1 - \cos\theta\right) \ . \tag{41}$$

The quantity h/mc is identified as the Compton wavelength. It's not exactly the size of the target, but is rather the effective 'profile' from the perspective of a photon.

Antiproton Production

Modern physics has forced us to discard the 'billiard ball' interpretation of partcile collisions - especially in the genesis of antiparticles. An experiment was conducted at Berkely that forced collisions among protons. Above a certain energy threshold, the reaction $p + p \rightarrow p + p + p + \bar{p}$ was observed, where the symbol $\bar{p}$ denotes an antiproton. In other words, *collisions among ordinary particles can create antimatter particles.*

Assume a setup where a speeding proton of energy E_0 and momentum $\vec{k}_0$ collides with a proton that is stationary in the laboratory frame. In four-vector notation, we may write this reaction as

$$P_1 + P_2 = P_A + P_B + P_C + \bar{P} \ . \tag{42}$$

Question: What is the minimum value of E_0 that allows this reaction to occur? Answer: we will evaluate the square of (42) in two different reference frames. This is allowed because the square of a four-vector is an invariant. *So long as we compute the square of a four-vector, the result is the same in all frames.*

Evaluate the left side of (42) in the laboratory frame. Explicitly,

$$\left(P_1 + P_2\right)^2 = \left[\left(\frac{E_0}{c}, \vec{k}_0\right) + (mc, 0)\right]^2 = -\left(\frac{E_0}{c} + mc\right)^2 + k_0^2 \ .$$

The momentum term k_0^2 may be eliminated by $E_0^2 = k_0^2c^2 + m^2c^4$. Evaluating the right side of (42) in the laboratory frame is difficult. We therefore choose the center-of-mass (abbreviated COM) reference frame, in where each of the $3 + 1$ protons are *stationary.* This is how we determine the minimal energy for the reaction - the products should have no extra kinetic energy, which corresponds to zero speed in the COM frame. Therefore,

$$\left(P_a + P_B + P_C + \bar{P}\right)^2 = \left[(4mc, 0)\right]^2 = -4^2m^2c^2 \ .$$

Equating the two above results yields an equation for E_0, and comes to

$$E_0 = 7mc^2 \,,$$

which is quite a price to pay when trying to produce only $2mc^2$ in new particles.

It's more efficient to design a machine that takes protons, each having energy E_0, and collides them head-on such that the laboratory frame is equivalent to the COM frame. In this case, the energy required is

$$E_0 = 2mc^2 \,. \tag{43}$$

Problem 28: Prove equation (43).

4 Tensors and Index Notation

4.1 Covariant and Contravariant Vectors

The basic object for representing events in special relativity is the four-vector; the 'square' of which is invariant for all reference frames. A four-vector is represented by a letter and an index, and the placement of that index has significance.

The position four-vector x^μ is called a *contravariant* vector, or also a type $(1,0)$ tensor. On the other hand, the vector x_μ is called *covariant*, and is also a type $(0,1)$ tensor. In general, x^μ and x_μ are *not* equivalent.

Manipulation of four-vectors is not done with matrices, but using tensors instead. A *tensor* is a generalized vector or generalized matrix - it's a cluster of numbers that may be 'viewed from different perspectives to see different faces', metaphorically speaking. The simplest four-vector manipulation by tensor is the Lorentz transformation, which simply coverts coordinates between inertial reference frames. The most general Lorentz transformation is an extension of equation (7), and reads

$$(x')^\mu = \sum_{\nu=0}^{3} \Lambda^\mu_\nu x^\nu, \tag{44}$$

where each of the Λ^μ_ν specify the coordinate transformation from x^μ to $(x')^\mu$.

Notice in equation (44) that the summed-over index ν appears twice - as a lower index on Λ and an upper index on x. For now on we will follow the *Einstein summation convention*, where the $\sum$ symbol is omitted. The rule is whenever we see a repeated index in the 'up' and 'down' positions, sum over that index from 0 to 3. Equation (44) shortens to

$$(x')^\mu = \Lambda^\mu_\nu x^\nu \,.$$

4.2 Metric Tensor

The Minkowski dot product is a four-dimensional version of the ordinary dot product, except the 'time' component has a minis sign. From now on I'll stop referring to Minkowski's dot product and instead point to the *flat metric* tensor. In flat space, the metric keeps track of the '+' and '−' signs on the square of the spacetime interval

$$dS^2 = -c^2 dt^2 + d\vec{x}^2 \,. \tag{45}$$

The flat metric tensor is written as

$$\eta_{\mu\nu} = \begin{bmatrix} \eta_{00} & \eta_{01} & \eta_{02} & \eta_{03} \\ \eta_{10} & \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{20} & \eta_{21} & \eta_{22} & \eta_{23} \\ \eta_{30} & \eta_{31} & \eta_{32} & \eta_{33} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{bmatrix} , \tag{46}$$

which is a type $(0, 2)$ symmetric tensor. A tensor is generally *not* a matrix; it may become matrix-like when written in $(1, 1)$ form. There also exists the *inverse metric*, a type $(2, 0)$ tensor, defined such that

$$\eta^{\mu\alpha}\eta_{\alpha\rho} = \delta^{\mu}_{\rho} , \tag{47}$$

where δ^{μ}_{ρ} is the Kroneker delta symbol, which is zero if $\mu \neq \rho$ and equal to 1 if $\mu = \rho$. It happens that the inverse flat metric is numerically equal to the flat metric, which is only true in special relativity.

4.3 Raising and Lowering Indices

Of the many uses of the metric tensor, one of its main operations is to change the (vertical) position of an index on some tensor. To convert x^{α} into x_{ν}, write

$$x_{\nu} = \eta_{\alpha\nu}x^{\alpha} ,$$

and working backwards we have

$$x^{\mu} = \eta^{\mu\rho}x_{\rho} .$$

For a more general example, you could encounter statements like

$$T^{\alpha\beta}_{\gamma\delta} = \eta_{\rho\gamma}T^{\alpha\beta\rho}_{\delta} .$$

4.4 Contraction

Contraction is the operation of equating one up-index and one down-index on a tensor or tensor product, and then summing over that index. For example, to transform the type $(2, 1)$ tensor $T^{\alpha\beta}_{\gamma}$ into a type $(1, 0)$ tensor, let $\beta = \gamma$, which by the Einstein summation convention, results in the new object T^{α}.

Contraction cannot be performed on a stand-alone four-vector; the simplest choice is to use the 'square'. The square of a four-vector is formally called the *norm*, and is invariant for all reference frames. The norm of the differential position four-vector is $dx^{\alpha}dx_{\alpha}$, which expands into

$$dx^{\alpha}dx_{\alpha} = \eta_{\mu\nu}dx^{\mu}dx^{\nu} = dS^2 , \tag{48}$$

the familiar 'spacetime interval' from equation (36). Notice the η_{00} term has stuck a minus sign on the time component as expected. Contraction formally replaces the Minkowski dot product.

Digging further back, equation (18) reminds us that the spacetime interval dS^2 may be replaced by $-c^2d\tau^2$, where τ is the proper time, giving

$$-c^2d\tau^2 = \eta_{\mu\nu}dx^{\mu}dx^{\nu} , \tag{49}$$

embodying essentially all of our notational achievements made thus far. It is worth memorizing!

Problem 29: Consider the four-vectors U and V. Prove:

$$U_\alpha V^\alpha = U^\mu V_\mu \tag{50}$$

Problem 30: Using contraction, show that the norm of the velocity four-vector U^μ is always equal to $-c^2$. In other words, prove:

$$U^\nu U_\nu = -c^2 \tag{51}$$

Problem 31: Consider a worldline $x^\mu(\lambda)$, parameterized as

$$x^0(\lambda) = q\lambda \qquad x^1(\lambda) = \frac{q}{2}\cos(\lambda)$$

$$x^2(\lambda) = \frac{q}{2}\sin(\lambda) \qquad x^4(\lambda) = 0\,,$$

where q is an arbitrary length constant. Compute the proper time τ between $\lambda = 0$ and $\lambda = 2\pi$. Answer:

$$\tau = \int \frac{1}{c}\sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu} = \frac{q}{c}\sqrt{1-\frac{1}{4}}\int_0^{2\pi} d\lambda = \frac{q}{c}\sqrt{3}\pi$$

Problem 32: Define the *four-acceleration* $a^\mu = dU^\mu/d\tau$, and also the *four-force* as

$$f^\mu = m\frac{dU^\mu}{d\tau}\,. \tag{52}$$

Take a τ-derivative of $U^\mu U_\mu$ to show that $f_\mu U^\mu = 0$, and also show that, if $\vec{F}$ is the usual 'three-force' vector, then:

$$f^\mu = \gamma\left(\frac{\vec{F}\cdot\vec{v}}{c}, \vec{F}\right) \tag{53}$$

Generalized Metric

The metric tensor is most generally denoted $g^{\mu\nu}$, and does far more than keep track of '+' and '−' signs. Notice that the information filling flat metric tensor (46) was plucked off of the spacetime interval $dS^2 = -c^2dt^2 + d\vec{x}^2$. This idea can be generalized for situations far beyond flat spacetime, where we generally would have

$$dS^2 = dx_\alpha dx^\alpha = g_{\mu\nu}dq^\mu dq^\nu$$

such that $g_{\mu\nu}$ is the metric for a curved space. The coordinates q^μ called *generalized coordinates* - their dimensionality is not necessarily length.

4.5 Lorentz Transformation Tensor

Given any reference frame, the sum of all (constant-velocity) boosts relative to that frame constitute only part of the 'family' of allowed Lorentz transformations. Reference frames that have been rotated may be trivially added to the list. The sum of such transformations is called f the*Lorentz group*. Members of the Lorentz group, most generally, obey the tensor equation

$$x'^{\alpha} = \Lambda^{\alpha}_{\mu} x^{\mu} \ .$$

Another change in reference frame is a translation in spacetime. Adding translations to the list of boosts and rotations, we construct the *Poincare* group

$$x'^{\alpha} = \Lambda^{\alpha}_{\mu} x^{\mu} + \tilde{x}^{\alpha} \ ,$$

which has 10 generations: 3 boosts , 3 fsrotations, and 4 translations.

We can write down a restriction on the components of Λ by equating the (invariant differential) spacetime interval as measured in separate inertial frames as follows:

$$\eta_{\alpha\beta} dx^{\alpha} dx^{\beta} = \eta_{\mu\nu} \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} dx^{\alpha} dx^{\beta}$$

Notice that the term $dx^{\alpha} dx^{\beta}$ is present on both sides, so we evidently have

$$\eta_{\alpha\beta} = \eta_{\mu\nu} \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} \ . \tag{54}$$

By various choices of α and β in (54), we interrelate certain components of Λ. For $\alpha = 0, \beta = 0$, equation (54) boils down to

$$1 = \left(\Lambda^{0}_{0}\right)^{2} - \left(\Lambda^{1}_{0}\right)^{2} \ ,$$

and for $\alpha = 1, \beta = 1$, we get

$$1 = \left(\Lambda^{1}_{1}\right)^{2} - \left(\Lambda^{0}_{1}\right)^{2} ,$$

and so on.

When the components of Λ represent a boost in the x-direction, we have found

$$\Lambda^{0}_{0} = \Lambda^{1}_{1} = \gamma \qquad \Lambda^{0}_{1} = \Lambda^{1}_{0} = -\gamma\beta \qquad \Lambda^{3}_{3} = \Lambda^{4}_{4} = 1 \ ,$$

with all other components equal to zero. The components of Λ change when boosting in another direction.

Problem 33: For a boost along the x-direction, show that equation (54) is satisfied by

$$\Lambda^{0}_{0} = \Lambda^{1}_{1} = \cosh(\phi) \qquad \Lambda^{1}_{0} = \Lambda^{0}_{1} = -\sinh(\phi) \ ,$$

where ϕ is the rapidity of a different inertial frame. Also, reproduce the Lorentz transformation equations (5) and (6) from the tensor Λ.

Problem 34: For two consecutive Lorentz boosts in the x-direction, show that the effective transformation tensor $\tilde{\Lambda}\Lambda$ in the expression

$$\left(x''\right)^{\mu} = \tilde{\Lambda}^{\mu}_{\alpha} \Lambda^{\alpha}_{\rho} x^{\rho}$$

reproduces the matrix in equation (15).

Inverse Lorentz Transformation

The *inverse Lorentz transformation* is the operation that 'undoes' the Lorentz transformation such that

$$x^{\mu} = \left(\Lambda^{-1}\right)^{\mu}_{\nu} \left(x'\right)^{\nu} \ ,$$

readily giving

$$\Lambda^{\mu}_{\nu} \left(\Lambda^{-1}\right)^{\nu}_{\gamma} = \delta^{\mu}_{\gamma} \ . \tag{55}$$

Coordinate Transformations in General

The symbol Λ^{α}_{ν} is reserved for coordinate transformations between inertial reference frames. In the general case, the Lorentz transformation of some object A^{μ} is written in terms of partial derivatives:

$$A'^{\alpha} = \frac{\partial x'^{\alpha}}{\partial x^{\mu}} A^{\mu} \tag{56}$$

4.6 Tensor Symmetry

Symmetric Tensors

A tensor is *symmetric* in two of its indices when it obeys

$$A^{\mu\nu} = A^{\nu\mu} \ ,$$

where A itself may itself be an object with more indices. In such a case, one may write

$$A^{\mu\nu} = \frac{1}{2}\left(A^{\mu\nu} + A^{\nu\mu}\right) + \frac{1}{2}\left(A^{\mu\nu} - A^{\nu\mu}\right) \ .$$

The symmetric part of a tensor is denoted using parentheses as

$$A^{(\mu\nu)} = \frac{1}{2}\left(A_{\mu\nu} + A^{\nu\mu}\right) \ . \tag{57}$$

Antisymmetric Tensors

A tensor is *antisymmetric* in two of its indices when it obeys

$$A^{\mu\nu} = -A^{\nu\mu} \ ,$$

where A itself may itself be an object with more indices. In such a case, one may (again) write

$$A^{\mu\nu} = \frac{1}{2}\left(A^{\mu\nu} + A^{\nu\mu}\right) + \frac{1}{2}\left(A^{\mu\nu} - A^{\nu\mu}\right) \ .$$

The antisymmetric part of a tensor is denoted using brackets as

$$A^{[\mu\nu]} = \frac{1}{2}\left(A_{\mu\nu} - A^{\nu\mu}\right) \ . \tag{58}$$

Corollary

Any tensor may be decomposed into symmetric and antisymmetric parts as follows:

$$A^{\mu\nu} = A^{(\mu\nu)} + A^{[\mu\nu]}$$

Problem 35: Consider the tensor $X^{\mu\nu}$ and the four-vector V^μ, given by

$$X^{\mu\nu} = \begin{bmatrix} 2 & 0 & 1 & -1 \\ -1 & 0 & 3 & 2 \\ -1 & 1 & 0 & 0 \\ -2 & 1 & 1 & -2 \end{bmatrix} \qquad\qquad V^\mu = (-1, 2, 0, -2) \ ,$$

respectively. Find the following quantities (i) X^μ_ν, (ii) X_μ^ν, (iii) $X^{(\mu\nu)}$, (iv) $X_{[\mu\nu]}$, (v) X^λ_λ, (vi) $V^\mu V_\mu$, (vii) $V_\mu X^{\mu\nu}$. Answer:

$$X^\mu_\nu = \begin{bmatrix} -2 & 0 & 1 & -1 \\ 1 & 0 & 3 & 2 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & -2 \end{bmatrix} \qquad\qquad X_\mu^\nu = \begin{bmatrix} -2 & 0 & -1 & 1 \\ -1 & 0 & 3 & 2 \\ -1 & 1 & 0 & 0 \\ -2 & 1 & 1 & -2 \end{bmatrix}$$

$$X^{(\mu\nu)} = \begin{bmatrix} 2 & -1/2 & 0 & -3/2 \\ -1/2 & 0 & 2 & 3/2 \\ 0 & 2 & 0 & 1/2 \\ -3/2 & 3/2 & 1/2 & -2 \end{bmatrix} \qquad\qquad X_{[\mu\nu]} = \begin{bmatrix} 0 & -1/2 & -1 & -1/2 \\ 1/2 & 0 & 1 & 1/2 \\ 1 & -1 & 0 & -1/2 \\ 1/2 & -1/2 & 1/2 & 0 \end{bmatrix}$$

$$X^\lambda_\lambda = -4 \qquad\qquad V^\mu V_\mu = 7 \qquad\qquad V_\mu X^{\mu\nu} = (4, -2, 5, 7)$$

5 Classical Field Theory

5.1 Fields and Coordinate Transformations

Vector Fields

A *field*, strictly speaking, is any function of coordinates x^μ. A *vector field* is a four-component vector that exists at all points in spacetime, represented as $V^\mu(x^\mu)$. In special relativity, the components of a vector field undergo a contravariant Lorentz transformation via

$$V^\mu\left(x^\alpha\right) \qquad \to \qquad V'^\mu\left(x'^\alpha\right) = \Lambda^\mu_\beta V^\beta\left(x^\alpha\left(x'^\alpha\right)\right) \tag{59}$$

in order to qualify as a field at all. Note the components of the vector V^μ change under Λ, but so also do the coordinates x^μ.

Problem 36: Take the vector field $V^\mu\left(x^\alpha\right) = (2x, 0, ct, 0)$ and determine $V'^\mu(x'^\alpha)$ if the primed reference frame is Lorentz-boosted along x. Answer:

$$(V')^\mu = \left[2\gamma^2\left(x' + vt'\right), -2\gamma^2\beta\left(x' + vt'\right), \gamma\left(ct' + vx'/c\right), 0\right]$$

Problem 37: The 'covariant Lorentz transformation' involves transforming the covariant vector field V_μ. Using equations (54) and (55), show that

$$\left(V'\right)_\mu = \left(\Lambda^{-1}\right)^\rho_\mu V_\rho \, . \tag{60}$$

The idea behind the Lorentz transformation extends beyond flat space. Taking the differential version of equation (44), we write

$$\frac{dx'^{\alpha}}{dx^{\mu}} = \Lambda^{\alpha}_{\mu} ,$$

which you may already know as the *Jacobian.* (This construction was previously proposed as equation (56).) In general then, for two reference frames x^{μ} and $\tilde{x}^{\mu}$, some vector field (four-vector) A_{μ} transforms as

$$A'_{\mu} = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} A_{\alpha} . \tag{61}$$

If it's difficult to remember which x-symbol has the prime ($'$) mark, put the prime symbol with the index instead of the letter, as in

$$A_{\mu'} = \frac{\partial x^{\nu}}{\partial x^{\mu'}} A_{\nu} .$$

For completeness, the contravariant version of (61) reads

$$A^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^{\alpha}} A^{\alpha} . \tag{62}$$

Scalar Fields

A *scalar field*, represented by $\phi(x^{\mu})$, is a function of coordinates whose value does not change under a coordinate transformation. For example, the temperature in a room is a scalar field - it's value at a point doesn't care whether we locate that point in Cartesian coordinates or in spherical coordinates.

<u>Problem 38</u>: Suppose there exists a scalar field $\phi(t, x, y, z) = t - 2y$. Express the scalar field ϕ as observed in an x-direction Lorentz-boosted frame. Answer: $\phi(t', x', y', z') = \gamma(t' + vx'/c^2) - 2y'$

It's obvious that a scalar should always transform into another scalar. To prove this though, consider the contraction of two four-vectors, namely $V^{\mu}U_{\mu}$. In another reference frame, the transformed quantity reads

$$\tilde{V}^{\mu}\tilde{U}_{\mu} = \frac{\partial \tilde{x}^{\mu}}{\partial x^{\alpha}} \frac{\partial x^{\beta}}{\partial \tilde{x}^{\mu}} V^{\alpha} U_{\beta} = \delta^{\beta}_{\alpha} V^{\alpha} U_{\beta} = V^{\beta} U_{\beta} ,$$

and is scalar as expected. Moreover, the contraction $V^{\mu}U_{\mu}$ is invariant.

Tensor Fields

The transformation rule for tensors is identical to that of vector fields. For example, equations (61) and (62) tell us any object with an 'up'-index shall Lorentz-transform using Λ, and for a 'down'-index we use the inverse, Λ^{-1}. Indices need not be transformed one at a time; for example one may legally write

$$\tilde{W}^{\alpha\beta}_{\gamma\rho} = \Lambda^{\alpha}_{\delta} \Lambda^{\beta}_{\epsilon} \left(\Lambda^{-1}\right)^{\nu}_{\gamma} W^{\delta\epsilon}_{\nu\rho} .$$

Index notation is necessary for expressing more general transformations. For example, two tensors $T^{\mu\nu}$ and S^{μ}_{ν} transform as:

$$\tilde{T}^{\rho\sigma}\left(\tilde{x}^{\beta}\right) = \frac{\partial \tilde{x}^{\rho}}{\partial x^{\mu}}\frac{\partial \tilde{x}^{\sigma}}{\partial x^{\nu}} T^{\mu\nu}(x^{\alpha}(\tilde{x}^{\beta}))$$

$$\tilde{S}^{\rho}_{\sigma}\left(\tilde{x}^{\beta}\right) = \frac{\partial \tilde{x}^{\rho}}{\partial x^{\mu}}\frac{\partial x^{\nu}}{\partial \tilde{x}^{\sigma}} S^{\mu}_{\nu}(x^{\alpha}(\tilde{x}^{\beta}))$$

5.2 Div, Grad, Curl in Index Notation

Four-Gradient

Consider a scalar field $\phi(x^{\mu})$. Write down the *four-gradient* components as

$$\frac{\partial \phi\left(x^{\mu}\right)}{\partial x^{\mu}} = \partial_{\mu}\phi = \left(\frac{\partial \phi}{\partial x^{0}}, \frac{\partial \phi}{\partial x^{1}}, \frac{\partial \phi}{\partial x^{2}}, \frac{\partial \phi}{\partial x^{3}}\right), \tag{63}$$

which is a covariant (low index) vector field. Under the Lorentz transformation $x^{\nu} = (\Lambda^{-1})^{\nu}_{\mu}(x')^{\mu}$, the four-gradient transforms as

$$\frac{\partial \phi}{\partial\left(x'\right)^{\mu}} = \frac{\partial \phi}{\partial x^{\nu}}\frac{\partial x^{\nu}}{\partial\left(x'\right)^{\mu}} = \left(\Lambda^{-1}\right)^{\nu}_{\mu}\partial_{\nu}\phi\,,$$

or in the general case,

$$\partial_{\mu'}\phi = \frac{\partial x^{\alpha}}{\partial x^{\mu'}}\partial_{\alpha}\phi\,.$$

The 'contravariant' four-gradient of ϕ reads

$$\partial^{\mu}\phi = \eta^{\mu\nu}\partial_{\nu}\phi\,, \tag{64}$$

which demonstrates a more general point: four-gradients can be used on a scalar field to create an object of any rank. For instance, $\partial_{\mu}\partial_{\nu}\phi$ is a type $(0,2)$ object. However, it would be a mistake to call these new objects tensors without checking their behavior under coordinate transformations.

Problem 39: Consider the covariant vector field $V_{\mu}(x^{\mu})$. Under general coordinate transformations $x^{\mu} \to \tilde{x}^{\mu}$, show that the quantity $\partial\left[_{\alpha}V_{\beta}\right]$ transforms as a type $(0,2)$ tensor. Answer:

$$\tilde{\partial}\left[_{\alpha}\tilde{V}_{\beta}\right] = \frac{1}{2}\left(\frac{\partial x^{\nu}}{\partial \tilde{x}^{\alpha}}\frac{\partial x^{\gamma}}{\partial \tilde{x}^{\beta}}\frac{\partial V_{\gamma}}{\partial x^{\nu}} - \frac{\partial x^{\nu}}{\partial \tilde{x}^{\beta}}\frac{\partial x^{\gamma}}{\partial \tilde{x}^{\alpha}}\frac{\partial V_{\gamma}}{\partial x^{\nu}}\right)$$

Four-Divergence

The *four-divergence* is an operation applied to vector fields (or objects with greater than one index):

$$\frac{\partial A^{\mu}}{\partial x^{\mu}} = \partial_{\mu}A^{\mu} \tag{65}$$

Much more will be said about the four-divergence later, but for now just consider equation (65) to be the generalization of $\vec{\nabla}\cdot\vec{A}$ for flat space.

Generalized Curl

The curl operator ($\vec{\nabla}\times$) does not have a simple four-dimensional analog. (Research the 'wedge product' for the next-best thing. An effective formula for the four-curl shall be written when we cover electromagnetism.) The cross product however is expressed conveniently with index notation by introducing the *Levi-Civita* symbol ϵ_{ijk}. If the indices i, j, k are an even permutation of the sequence $1, 2, 3$, then $\epsilon = +1$. For odd permutations we have $\epsilon = -1$. If any two indices are equal, $\epsilon = 0$. To generalize to beyond three dimensions, add more indices to ϵ. The cross product is therefore:

$$\left(\vec{a} \times \vec{b}\right)_i = \epsilon_{ijk} a^j b^k \tag{66}$$

With equation (66), the three-dimensional curl may be written as

$$\left(\vec{\nabla} \times \vec{F}\right)^i (\vec{x}) = \epsilon^{ijk} \frac{\partial}{\partial x^j} F^k (\vec{x}) \ . \tag{67}$$

5.3 Relativistic Lagrangian

In nonrelativistic physics, equations of motion are obtained from the Lagrangian $L(\vec{x}, \dot{\vec{x}}, t)$. Defining the *action* $S = \int L dt$, and then enforcing the *principle of least action*, from which the Euler-Lagrange equations emerge.

$$\delta \int L dt = 0 \, . \tag{68}$$

For instance, suppose the Lagrangian for a particle is a function of the position $q(t)$ and velocity $\dot{q}(t)$. Extremizing the action $S = \int L dt$, write

$$\delta S = \int \left[\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right] dt \, .$$

Integrating by parts and setting the boundary terms to zero, we get

$$\delta S = \int \delta q \left[\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) \right] dt \, .$$

Setting the left side to zero, pick out the the nonrelativistic Euler-Lagrange equation:

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) = \frac{\partial L}{\partial q} \tag{69}$$

Lagrangian mechanics must be redesigned to account for special relativity, namely because spacetime coordinates depend on the frame of reference. It makes sense to recast S in terms of an invariant with respect to reference frame. A natural choice therefore is to integrate the proper time of the particle (including a conventional minus sign and adjusting for units):

$$S = -mc^2 \int d\tau = -mc^2 \int dt \sqrt{1 - \frac{u^2}{c^2}} \tag{70}$$

In the last step, $d\tau$ has been replaced by dt/γ in accordance with equation (3). Evidently, the kinetic component of the Lagrangian is not $mu^2/2$, but is instead $-\sqrt{1 - u^2/c^2}$. In the low-velocity limit, we indeed find

$$-mc^2 \sqrt{1 - \frac{u^2}{c^2}} \approx -mc^2 + \frac{mu^2}{2} \, ,$$

where the rest mass is correctly subtracted away, and the kinetic term has reduced to its Newtonian limit.

Another way to derive the relativistic Lagrangian is to begin with the canonical definition of the momentum of a particle,

$$p_i = \frac{\partial L}{\partial u_i} \,,$$

and then substitute the relativistic momentum found previously, namely $p_i = \gamma m u_i$. Intrgrating, we find (tacking on the potential term)

$$L = -mc^2\sqrt{1-\frac{u^2}{c^2}} - U\left(\vec{r}\right) \,. \tag{71}$$

5.4 Hamiltonian Formalism

The classical Hamiltonian is defined by the Legendre transformation:

$$H\left(q_1, q_2, \ldots, p_1, p_2, \ldots\right) = \sum_j p_j \dot{q}_j - L\left(q_1, q_2, \ldots, \dot{q}_1, \dot{q}_2, \ldots\right) \,, \tag{72}$$

where L is the Lagrangian and the sum is over all j coordinates. Note that each p_j, q_j, and $\dot{q}_j$ is a function of time, and the dot ($\dot{}$) symbol over a variable denotes a time derivative. In shorthand notation, we'll write

$$H\left(\vec{q}, \vec{p}\right) = \sum p_j \dot{q}_j - L\left(\vec{q}, \dot{\vec{q}}\right) \,.$$

The *canonical momentum* p_j is defined from the Lagrangian as

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \,. \tag{73}$$

In the special case where the Lagrangian takes the form $L = \sum m\dot{q}_j^2/2 - V(\vec{q})$, the Hamiltonian reduces to

$$H = \sum_j \left(\dot{q}_j p_j - \frac{1}{2} m \dot{q}_j^2\right) + V\left(\vec{q}\right) \,.$$

If we make the assumption $p = m\dot{q}$, then we recover the formula for the total energy of the system:

$$E = H = \sum_j \frac{p_j^2}{2m} + V\left(\vec{q}\right)$$

Returning to the general case with L unspecified, minimize the action $S = \int L dt$. Replacing L by equation (72), we write

$$\delta S = \delta \int \left[\sum_j p_j \dot{q}_j - H\left(\vec{q}, \vec{p}\right)\right] dt \,,$$

which to linear order in p_j, q_j, and $\dot{q}_j$, becomes

$$\delta S = \delta \int \sum_j \left[\delta p_j \dot{q}_j + p_j \delta \dot{q}_j - \delta p_j \frac{\partial H}{\partial p_j} - \delta q_j \frac{\partial H}{\partial q_j}\right] dt \,.$$

Letting the left side equal to zero, we discover the pair of relations

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad\qquad \dot{p}_j = -\frac{\partial H}{\partial q_j}\,, \tag{74}$$

known as *Hamilton's equations* of motion.

Poisson Brackets

The *Poisson bracket* is an operation on two differentiable functions $A(\vec{q},\vec{p})$ and $B(\vec{q},\vec{p})$ such that

$$\{A,B\}_{qp} = \sum_j \left(\frac{\partial A}{\partial q_j}\frac{\partial B}{\partial p_j} - \frac{\partial A}{\partial p_j}\frac{\partial B}{\partial q_j}\right)\,. \tag{75}$$

Right away, we can see two properties that always hold:

$$\{A,B\}_{qp} = -\{B,A\}_{qp} \qquad\qquad \{A,A\}_{qp} = 0 \tag{76}$$

Consider some function $F(\vec{q},\vec{p},t)$. By the chain rule, we know

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \sum_j \left(\frac{\partial F}{\partial q_j}\frac{dq_j}{dt} + \frac{\partial F}{\partial p_j}\frac{dp_j}{dt}\right)\,,$$

which, after using equations (74) and (75), becomes

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \{F,H\}_{qp}\,. \tag{77}$$

If the system is isolated from its environment, the total energy E must be conserved, making $\partial H/\partial t = 0$. Poisson brackets immediately hand us the stronger statement that $dH/dt = 0$ and $E = H$. It follows that the Hamiltonian can be used as the generator of infinitesimal time translations.

Problem 40: Using equation (75), prove the relations:

$$\dot{q}_k = \{q_k, H\}_{qp} \qquad\qquad \dot{p}_k = \{p_k, H\}_{qp} \qquad\qquad \{q_j, p_k\}_{qp} = \delta_{jk} \tag{78}$$

5.5 Euler-Lagrange Equation for Fields

Classical field theory shares the same philosophical starting point as Lagrangian mechanics, that is, the principle of least action. To generalize Lagrange and Hamiltonian formaism for fields, the bold innovation is that generalized coordinates $q(t)$ are replaced by spacetime-dependent fields $\phi(\vec{x},t)$, also written $\phi^\mu(x^\mu)$, along with their derivatives. The Lagrangian L shall be recast in terms of a Lorentz scalar known as the Lagrangian density $\mathcal{L}$, as

$$L = \int \mathcal{L}\left(\phi^\mu, \partial_\alpha \phi^\mu\right) d^3x\,.$$

The action S is defined as

$$S = \int L dt\,,$$

so correspondingly

$$S = \int \mathcal{L}\left(\phi^\mu, \partial_\alpha \phi^\mu\right) d^4x\,.$$

Minimizing the action S leads to a set of Euler-Lagrange equations. To proceed, write the field(s) ϕ^μ under small variations,

$$\phi^\mu \to \phi^\mu + \delta\phi^\mu$$
$$\partial_\alpha \phi^\mu \to \partial_\alpha \phi^\mu + \partial_\alpha \left(\delta\phi^\mu\right) \, .$$

along with the first-order Taylor expansion of the Lagrangian density

$$\mathcal{L}\left(\phi^\mu, \partial_\alpha \phi^\mu\right) \to \mathcal{L}\left(\phi^\mu, \partial_\alpha \phi^\mu\right) + \frac{\partial \mathcal{L}}{\partial \phi^\mu}\delta\phi^\mu + \frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \phi^\mu\right)}\partial_\alpha\left(\delta\phi^\mu\right) \, ,$$

such that the variation in the action is

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial \phi^\mu}\delta\phi^\mu + \frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \phi^\mu\right)}\partial_\alpha\left(\delta\phi^\mu\right)\right] d^4x \, .$$

Assume for now that all variations are zero on the integration boundaries. The second term in δS must be integrated by parts, specifically with

$$u = \frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \phi^\mu\right)} \qquad\qquad dv = \partial_\alpha\left(\delta\phi^\mu\right) d^4x \, ,$$

giving

$$\delta S = \int \delta\phi^\mu \left[\frac{\partial \mathcal{L}}{\partial \phi^\mu} - \partial_\alpha\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\alpha \phi^\mu\right)}\right)\right] d^4x \, .$$

Applying the principle of least action, we let $\delta S = 0$, and it follows that the quantity in square brackets must equal zero. We have found the *Euler-Lagrange equation for fields*:

$$\frac{\partial \mathcal{L}}{\partial \phi^\mu} - \frac{\partial}{\partial x_\alpha}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial \phi^\mu / \partial x_\alpha\right)}\right) = 0 \tag{79}$$

5.6 Scalar Field Theory

The simplest example of a classical field $\phi(x^\mu)$ is a real scalar field, in which case (as in all cases) the classical mechanics are derived from energy density considerations. We know loosely that energy has both kinetic and various potential forms, but in assembling the Lagrangian we must choose some combination of energy terms that is Lorentz-invariant. The potential energy density $U(\phi)$ is already Lorentz-invariant, but the 'usual' expressions for kinetic terms don't work (try a few).

Consider the Lorentz-invariant quantity $\tilde{\mathcal{L}}$ as written in flat space

$$\tilde{\mathcal{L}} = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi = \frac{1}{2c^2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 \, ,$$

which contains both a time-derivative term $(\dot{\phi})^2/2$ and a gradient energy term $(\nabla\phi)^2$. Interpreting $\tilde{\mathcal{L}}$ as the kinetic energy density, the proper Lagrangian density (in accordance with $L = T - U$) for a real scalar field is

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - U\left(\phi\right) \, . \tag{80}$$

Now apply the Euler-Lagrange equation (79) to the Lagrangian density (80). The term involving U is trivial, but the other is not:

$$\frac{\partial \mathcal{L}}{\partial \phi} = -\frac{dU}{d\phi} \qquad\qquad \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} = \frac{\partial}{\partial\left(\partial_\mu \phi\right)}\left[-\frac{1}{2}\eta^{\alpha\beta}\left(\partial_\alpha\phi\right)\left(\partial_\beta\phi\right)\right]$$

Note that the indices α, β have been introduced to prevent repetitive labeling. Since a change of labels can bring about a zoo of unwanted characters, I point out the following identity for the Kroneker delta symbol,

$$\frac{\partial A_\rho}{\partial A_\sigma} = \delta^\sigma_\rho \,,$$

where A stands for 'any' object. The $\partial_\mu\phi$-derivative of $\mathcal{L}$ is

$$\frac{\partial}{\partial\left(\partial_\mu\phi\right)} = -\frac{1}{2}\eta^{\alpha\beta}\delta^\mu_\alpha\partial_\beta\phi - \frac{1}{2}\eta^{\alpha\beta}\partial_\alpha\phi\delta^\mu_\beta = -\eta^{\mu\rho}\partial_\rho\phi \,,$$

and the Euler-Lagrange equation becomes

$$-\frac{dU}{d\phi} + \eta^{\mu\nu}\partial_\mu\partial_\nu\phi = 0 \,.$$

<u>Problem 41</u>: Consider the three scalar fields ϕ_i living in three-dimensional flat space, with $i = (x, y, z)$. Next, rotate the field ϕ_i about an arbitrary unit vector n_j by an infinitesimal angle $\Delta\theta$. Show that the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - \frac{1}{2}m^2\phi_i\phi_i$$

is invariant under the rotation

$$\phi_i \to \phi_i + \Delta\theta\epsilon_{ijk}n_j\phi_k \,.$$

In the language of group theory, you are to show that the $\mathcal{L}$ given has $SO\,(3)$ symmetry. Answer: Each term can be handled separately, and each term linear in $\Delta\theta$ is zero.

$$\partial_\mu\phi_i\partial^\mu\phi_i \to \partial_\mu\phi_i\partial^\mu\phi_i + 2\Delta\theta\epsilon_{ijk}\partial_\mu\phi_i\partial^\mu\phi_k = \partial_\mu\phi_i\partial^\mu\phi_i$$

$$\phi_i\phi_i \to \phi_i\phi_i + 2\Delta\theta\epsilon_{ijk}n_j\phi_i\phi_k = \phi_i\phi_i$$

D'Alemtertian Operator

The combination $\eta^{\mu\nu}\partial_\mu\partial_\nu$ has a special name called the *D'Alemtertian operator*, which is essentially the Minkowski generalization of the Laplacian operator $\nabla^2 = \Delta$. Because the D'Alembertian operator deals with four dimensions, it's symbol is the square, particularly

$$\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \nabla^2 \,.$$

The Euler-Lagrange equation for a real scalar field ϕ is therefore

$$\Box\phi - \frac{dU}{d\phi} = 0 \,. \tag{81}$$

(Note that the units of ϕ must be $\sqrt{Joule/meter}$.)

Klein-Gordon Equation

Perhaps the most common and useful form for $U(\phi)$ is the simple harmonic oscillator, namely

$$U(\phi) = \frac{1}{2}\tilde{m}^2\phi^2\,,$$

where $\tilde{m} = mc/\hbar$ is a constant that has units of inverse length, but is referred to as the 'mass' of the field. Recall that the combination h/mc is the Compton wavelength. The Euler-Lagrange equation becomes the *Klein-Gordon equation*

$$\Box\phi - \tilde{m}^2\phi = 0\,, \tag{82}$$

which is easily solved by plane waves, namely

$$\phi = \phi_0 e^{i(Et - \vec{p}\cdot\vec{x})/\hbar}\,. \tag{83}$$

The momentum $\vec{p}$ is related to p_0 through the energy-momentum relation (25), namely

$$E^2 = (\vec{p})^2 c^2 + m^2 c^4\,.$$

Hamiltonian Formalism for Scalar Fields

The notion of canonical momentum shall be generalized for scalar fields. To do this, we will interpret the field $\phi(x^\mu)$ as a set of *canonical coordinates*. Therefore, the canonical momenfum for fields may be written

$$\Pi(x^\mu) = \frac{\partial\mathcal{L}}{\partial\partial_0\phi(x^\mu)}\,, \tag{84}$$

also called the *momentum density*. Note the operator ∂_0 is equivalent to $c^{-1}\partial_t$. In the same way that $H = \sum p\dot{q} - L$, let us analogously define the *Hamiltonian density* as

$$\mathcal{H}(\phi, \Pi) = \Pi\partial_0\phi(x^\mu) - \mathcal{L}(\phi, \partial_\mu\phi) \tag{85}$$

Substituting the Lagrangian density (80), namely

$$\mathcal{L} = \frac{1}{2c^2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 - U(\phi)$$

into the above gives the Hamiltonian density for real scalar fields:

$$\mathcal{H} = \frac{1}{2}\left(\Pi(x^\mu)\right)^2 + \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 + U(\phi) \tag{86}$$

The notion of Poisson brackets must also apply to fields. Of course, the object Π is a density, so the generalization of equation (75) must involve an integral:

$$\{A, B\}_{\phi\Pi} = \int\left(\frac{\partial A}{\partial\phi(x^\mu)}\frac{\partial B}{\partial\Pi(x^\mu)} - \frac{\partial A}{\partial\Pi(x^\mu)}\frac{\partial B}{\partial\phi(x^\mu)}\right)d^3x \tag{87}$$

Note that A and B are two functionals that depend on ϕ and Π.

Finally, we can write an orthogonality condition analogous to equation (78), called the *equal-time Poisson bracket relation*:

$$\left\{\phi\left(x^0, \vec{x}\right), \Pi\left(x^0, \vec{y}\right)\right\}_{\phi\Pi} = \delta(\vec{x} - \vec{y}) \tag{88}$$

Connection to Classical Statistical Mechanics

The objects of classical field theory ϕ, Π can actually be used to construct the partition function Z, the main object of statistical mechanics. We shall use the notion of *imaginary time* by letting $t \to -it$, or in four-vector notation, $x^0 \to -ix^0$. The operator ∂_0 is replaced with $i\partial_0$.

Under this transformation, the action becomes

$$S_i = -i \int \mathcal{L}_i \left(\phi, i\partial_0\phi, \partial_j\phi\right) d^4x \ ,$$

where $\mathcal{L}_i$ is the Lagrangian density. Note that the operator ∂_j excludes the time component altogether. If the Lagrangian density takes a simple form like equation (80), the imaginary-time version reads

$$\mathcal{L}_i = -\frac{1}{2c^2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 - U\left(\phi\right) \ .$$

The action is therefore

$$S_i = i\int \left(\frac{1}{2c^2}\left(\frac{\partial\phi}{\partial t}\right)^2 + \frac{1}{2}\left(\vec{\nabla}\phi\right)^2 + U\left(\phi\right)\right) d^4x \ .$$

The action now extremizes some energy quantity that lives in $0+4$ dimensions instead of $1+3$. With time removed form the picture, the idea of extremizing potentials of the above form is the name of the game in statistical mechanics. Defining $\mathcal{D}\phi$ as the density of states, the partition function of a classical system reads

$$Z = \int \left(\mathcal{D}\phi\right) \exp\left[iS\left(\phi, i\partial_0\phi, \partial_j\phi\right)/\hbar\right] \ .$$

5.7 Noether's Theorem

Emmy Noether discovered a deep and fundamental propertiy of the Lagrangian that has a huge impact on theoretical physics. *Noether 's theorem* states that any differentiable symmetry of the action S of a physical system has a corresponding conservation law. It's not obvious from Newtonian mechanics that time translation symmetry corresponds to conservation of energy, rotational symmetry leads to conservation of angular momentum, and spatial translation symmetry gives rise to conservation of momentum. The power of Noether 's theorem is that *any* symmetry yields a conservation law!

Noether Current

Consider an infinitesimal change in the field $\phi(x^\mu)$ such that

$$\phi\left(x^\mu\right) \to \phi\left(x^\mu\right) + \delta\phi\left(x^\mu\right) \ .$$

In response to the change $\delta\phi$, the Lagrangian density behaves as $\mathcal{L} \to \mathcal{L} + \delta\mathcal{L}$, where the change $\delta\mathcal{L}$ shall be expressed as a four-divergence $\partial_\mu \mathcal{J}^\mu$ as

$$\delta\mathcal{L} = \partial_\mu \mathcal{J}^\mu \ .$$

Now take the variation of the action $S = \int d^4x\mathcal{L}(\phi, \partial_\mu\phi)$ in the region R. First-order Taylor expansion gives

$$\delta S = \int_R d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\partial_\mu\delta\phi\right] ,$$

where the equality $\delta\partial_\mu\phi = \partial_\mu\delta\phi$ has been used. Notice the above two terms are precisely those that led to the Euler-Lagrange equation for fields, (79), butf the boundary term that arises from integration by parts shall not be taken as zero. The variation in S reduces to

$$\delta S = \int_R d^4x \left[\partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\right)\delta\phi\right] = \int_R d^4x \left[\delta\mathcal{L}\right] ,$$

implying

$$\partial_\mu\left[\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\delta\phi - \mathcal{J}^\mu\right] = 0 .$$

We now define the *Noether current* $j^\mu(x^\mu)$ to be

$$j^\mu = \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu\phi\right)}\delta\phi - \mathcal{J}^\mu , \tag{89}$$

which is obviously conserved, obeying $\partial_\mu j^\mu = 0$.

Problem 42: Consider the three scalar fields ϕ_i living in three-dimensional flat space, with $i = (x, y, z)$. Calculate the Noether current j^μ associated with infinitesimal $SO(3)$ rotations, namely

$$\phi_i \to \phi_i + \Delta\theta\epsilon_{ijk}n_j\phi_k .$$

Answer:

$$j^\mu = \epsilon_{ijk}n_j\partial^\mu\phi_i\phi_k$$

Problem 43: Given the Noether current $j^\mu = \epsilon_{ijk}n_j\partial^\mu\phi_i\phi_k$, show that the charges

$$Q_i = \int j^0 d^3x$$

are conserved. Answer:

$$\dot{Q}_i = \int d^3x\epsilon_{ijk}n_j\left(\partial_{tt}\phi_i\phi_k + \dot{\phi}_i\dot{\phi}_k\right) = 0 .$$

5.8 Stress-Energy Tensor

Many field theories include the notion of the *stress-energy tensor*, also known as the *energy-momentum tensor*. The stress-energy tensor is defined as a two-index object written as $T^{\mu\nu}$, and contains (you guessed it) all information about the energy, momentum, pressure, etc, of some field ϕ.

The stress-energy tensor $T^{\mu\nu}$ is the conserved Noether current associated with spacetime translations. To see this, start with equation (89) and let

$$x^\mu \to x^\mu - a^\mu .$$

To first order, the field $\phi(x^\mu)$ responds to the change by

$$\delta\phi = -a^\nu\partial_\nu\phi ,$$

and the four-vector $\mathcal{J}^\mu$ is

$$\mathcal{J}^\mu = -a^\mu \mathcal{L}\,.$$

The Noether current becomes

$$j^\mu = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} a^\nu \partial_\nu \phi + a^\mu \mathcal{L} = a_\nu \left[-\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial^\nu \phi + \eta^{\mu\nu} \mathcal{L}\right]\,.$$

The term in square brackets is identified as the stress-energy tensor $\hat{T}^{\mu\nu}$. Explicitly, we have found

$$j^\mu = a_\nu \hat{T}^{\mu\nu} \qquad\qquad \hat{T}^{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial^\nu \phi + \eta^{\mu\nu} \mathcal{L}\,, \tag{90}$$

and because we know the Noether current is conserved, it follows that

$$\partial_\mu \hat{T}^{\mu\nu} = 0\,.$$

Note the 'hat' symbol (ˆ) above each letter T, which serves to foreshadow further refinements to $\hat{T}^{\mu\nu}$ based on symmetry arguments (see below).

Repeating the calculation from the ground, take the variation of the action $S = \int d^4x \mathcal{L}(x^\mu, \phi, \partial_\mu \phi)$ in the region R given the changes $x^\mu \to x^\mu - a^\mu$ and $\delta\phi = -a^\nu \partial_\nu \phi$. First-order Taylor expansion gives

$$\delta S = \int_R d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial_\mu \delta\phi + \frac{\partial \mathcal{L}}{\partial x^\mu} a^\mu\right]\,,$$

where the equality $\delta\partial_\mu \phi = \partial_\mu \delta\phi$ has been used, and recall δx^ν is equal to a^ν. From vector calculus, the final term in the above shall be recast as a boundary integral on ∂R. Note that the first two terms in the above lead to Euler-Lagrange equation (79) for fields, leaving the boundary integral left over from integration by parts. Letting all integrals over R cancel out, and meanwhile factoring a^ν out of the brackets gives

$$\delta S = \int_{\partial R} d_\mu \sigma a^\nu \left[-\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial_\nu \phi + \delta^\mu_\nu \mathcal{L}\right]\,.$$

From the outset, we are letting δS equal zero, and meanwhile a^ν is infinitesimal. The (nonzero) term in square brackets is none other than the stress energy tensor in $(1,1)$ form:

$$\hat{T}^\mu_\nu = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial_\nu \phi + \delta^\mu_\nu \mathcal{L} \tag{91}$$

With all raised indices, equation (91) becomes

$$\hat{T}^{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\right)} \partial^\nu \phi + \eta^{\mu\nu} \mathcal{L}\,. \tag{92}$$

Stress-Energy Tensor for Real Scalar Field

Consider the action

$$S = \int d^4x \sqrt{-\eta} \mathcal{L}\,, \tag{93}$$

where the index-free symbol η is the determinant of the metric tensor, and $\mathcal{L}$ is the Lagrangian density. For demonstration we shall substitute equation (80) for $\mathcal{L}$, the Lagrangian density of a real scalar field. The term that reads $\sqrt{-\eta}$ is the square root of the determinant of the metric. The inclusion of this term guarantees that d^4x Lorentz-transforms properly.

Next we vary the action with respect to small changes in the inverse metric, meaning

$$\eta^{\mu\nu} \to \eta^{\mu\nu} + \delta\eta^{\mu\nu} .$$

Before trying to tackle δS directly, do the grunt work first. From the statement

$$\delta^{\mu}_{\alpha} = \left(\eta^{\mu\rho} + \delta\eta^{\mu\rho}\right)\left(\eta_{\mu\sigma} + \delta\eta_{\mu\sigma}\right) ,$$

it follows that

$$\delta\eta_{\alpha\rho} = -\eta_{\alpha\mu}\eta_{\rho\nu}\delta\eta^{\mu\nu} .$$

Next, the variation of $\sqrt{-\eta}$ comes out to

$$\delta\sqrt{-\eta} = -\frac{1}{2\sqrt{-\eta}}\delta\eta = -\frac{1}{2}\sqrt{-\eta}\left(\eta_{\mu\nu}\delta\eta^{\mu\nu}\right) .$$

Using the Lagrangian density for a real scalar field, namely $\mathcal{L} = -\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi/2 - U(\phi)$, we also have

$$\delta\mathcal{L} = -\frac{1}{2}\delta\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$$

Now we may evaluate $\delta S = S(\eta + \delta\eta) - S(\eta)$ with ease, as

$$\delta S = \int d^4x \left[\mathcal{L}\delta\sqrt{-\eta} + \sqrt{-\eta}\delta\mathcal{L}\right] ,$$

simplifies to

$$\delta S = \int d^4x \left(-\frac{\sqrt{-\eta}}{2}\right)\delta\eta^{\mu\nu}\left[\partial_\mu\phi\partial_\nu\phi - \eta_{\mu\nu}\left(\frac{1}{2}\eta^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi + U(\phi)\right)\right] .$$

The quantity in square brackets is the stress-energy tensor with lowered indices, $T_{\mu\nu}$. With raised indices, the stress-energy tensor of a real scalar field ϕ is

$$T^{\mu\nu} = \eta^{\mu\lambda}\eta^{\nu\sigma}\partial_\lambda\phi\partial_\sigma\phi - \eta^{\mu\nu}\left[\frac{1}{2}\eta^{\lambda\sigma}\partial_\lambda\phi\partial_\sigma\phi + U(\phi)\right] , \tag{94}$$

a special case of (92), up to a conventional minus sign.

General Properties of Stress-Energy Tensor

The most important property of the stress-energy tensor $T^{\mu\nu}$ is that it has zero divergence, as in

$$\partial_\mu T^{\mu\nu} = 0 . \tag{95}$$

The stress-energy tensor should also (i) be symmetric, (ii) be traceless, and (iii) have T^{00} positive-definite. Respectively, these mean:

$$T^{\mu\nu} = T^{\nu\mu} \qquad T^{\alpha}_{\alpha} = 0 \qquad T^{00} > 0 \tag{96}$$

The problem with our most general stress-energy tensor, equation (92), is that $\hat{T}^{\mu\nu}$ is only symmetric when ϕ is a scalar field. When we consider vector fields ϕ^μ, the tensor $\mathrm{T}^{\mu\nu}$ is neither symmetric nor traceless. The remedy is to introduce a tensor $\mathcal{K}$ as a total divergence such that

$$T^{\mu\nu} = \hat{T}^{\mu\nu} + \partial_\lambda\mathcal{K}^{[\lambda\mu]\nu} , \tag{97}$$

where $\mathcal{K}$ is a three-index tensor that is antisymmetric in the indices λ, μ, meaning

$$\mathcal{K}^{[\lambda\mu]\nu} = \frac{1}{2}\left(\mathcal{K}^{\lambda\mu\nu} - \mathcal{K}^{\mu\lambda\nu}\right) .$$

Note we introduce $\mathcal{K}$ as a total divergence because $\partial_\mu\partial_\lambda\mathcal{K}^{[\lambda\mu]\nu} = 0$ must be true.

6 Electromagnetism

6.1 Maxwell's Equations

The fundamental objects of electromagnetism are the electric field $\vec{E}$ and the magnetic field $\vec{B}$. These fields vary in space and time according to Maxwell's equations (in SI units):

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \tag{98}$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{99}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{100}$$

$$\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \tag{101}$$

It turns out that Maxwell's equations describe all phenomena concerning electricity and magnetism. Continuity of electric charge is governed by

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0\,, \tag{102}$$

and is easily proven using the divergence theorem.

Of course, it's often easier to deal with vector or scalar potentials instead of the actual fields. From vector calculus we know:

$$\vec{B} = \vec{\nabla} \times \vec{A} \tag{103}$$

$$\vec{E} = -\vec{\nabla}\phi - \partial \vec{A}/\partial t \tag{104}$$

For completeness, I'll include that for a given charge distribution $\rho(\vec{x}, t)$ and current density $\vec{J}(\vec{x}, t)$, the scalar and vector potentials are given by

$$\phi\left(\vec{x}, t\right) = \frac{1}{4\pi\epsilon_0} \int d^3x' \frac{\rho\left(\vec{x}', t_r\right)}{|\vec{x} - \vec{x}'|} \tag{105}$$

$$\vec{A}\left(\vec{x}, t\right) = \frac{\mu_0}{4\pi} \int d^3x' \frac{\vec{J}\left(\vec{x}', t_r\right)}{|\vec{x} - \vec{x}'|}\,, \tag{106}$$

where

$$t_r = t - \frac{|\vec{x} - \vec{x}'|}{c} \tag{107}$$

is the *retarded time*.

Problem 44: Use Maxwell's equations to derive the electromagnetic wave equation(s) in the presence of no charges or currents:

$$-\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} + \nabla^2 \vec{E} = 0 \tag{108}$$

$$-\frac{1}{c^2}\frac{\partial^2 \vec{B}}{\partial t^2} + \nabla^2 \vec{B} = 0 \tag{109}$$

Problem 45: Use Maxwell's equations and the vector and scalar potentials to derive the equations of motion:

$$\nabla^2\phi + \frac{\partial\left(\vec{\nabla}\cdot\vec{A}\right)}{\partial t} = -\frac{\rho}{\epsilon_0} \tag{110}$$

$$\nabla^2\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = -\mu_0\vec{J} + \vec{\nabla}\left(\vec{\nabla}\cdot\vec{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right) \tag{111}$$

6.2 Electromagnetic Field Strength Tensor

Combine the scalar potential ϕ and the vector potential $\vec{A}$ into one object called the *electromagnetic four-potential*:

$$A^\alpha = \left(\phi/c, \vec{A}\right) \tag{112}$$

The fields $\vec{E}$ and $\vec{B}$ shall also be combined into one object. To proceed we use the Euler-Lagrange technique, but not by applying (71) directly. Instead the kinetic portion of L is taken to be $(m/2)U^\mu U_\mu$, and we also introduce a 'minimal coupling' to the electromagnetic potential A_μ such that the Lagrangian takes the form

$$L = \frac{m}{2}U^\mu U_\mu + qA_\mu U^\mu\ . \tag{113}$$

Cranking (113) through the Euler-Lagrange equation (69), and using the metric and the chain rule along the way, arrive at

$$m\frac{dU_\mu}{d\tau} + q\frac{dA_\mu}{d\tau} = q\frac{\partial A_\nu}{\partial x^\mu}U^\nu\ .$$

A bit more algebra delivers

$$m\frac{dU_\mu}{d\tau} = qU^\nu\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)\ ,$$

which has close resemblance to Newton's second law. The quantity in parenthesis is identified as the *electromagnetic field strength tensor*, having definition

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\ , \tag{114}$$

also written

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\ .$$

Observe that $F_{\mu\nu}$ is an antisymmetric tensor, which means

$$F_{\nu\mu} = -F_{\mu\nu}\ , \tag{115}$$

indicating that the 'diagonal' entries $F_{\alpha\alpha}$ are identically zero. There are only six independent components in F.

The nonzero components of $F_{\mu\nu}$ are straightforward to evaluate using the four-potential $A_\mu = (-\phi/c, \vec{A})$. Begin with F_{0i}, with $i = (1, 2, 3)$, giving the i-th component of the electric field:

$$F_{0i} = \left(\partial_t\vec{A}/c + \vec{\nabla}\phi/c\right)_i = -\frac{E_i}{c}$$

The magnetic field components are contained in terms F_{ij}, namely:

$$F_{23} = \left(\vec{\nabla} \times \vec{A}\right)_x \qquad F_{13} = -\left(\vec{\nabla} \times \vec{A}\right)_y \qquad F_{12} = \left(\vec{\nabla} \times \vec{A}\right)_z$$

Altogether we have, in block form (not a matrix):

$$F_{\mu\nu} = \begin{bmatrix} F_{00} & F_{01} & F_{02} & F_{03} \\ F_{10} & F_{11} & F_{12} & F_{13} \\ F_{20} & F_{21} & F_{22} & F_{23} \\ F_{30} & F_{31} & F_{32} & F_{33} \end{bmatrix} = \begin{bmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{bmatrix} \tag{116}$$

The components F_{ik} can be written more clearly in terms of $\vec{B}$ by using the Levi-Civita symbol:

$$F_{ik} = \sum_{k=1}^{3} \epsilon_{ijk} B_k$$

In summary, we have found (after ajdusting the indices for aesthetics), the relativistic equation of motion of a particle in an mixed electric and magnetic fields in flat space:

$$\frac{d^2 x^\mu}{d\tau^2} = \frac{q}{m} F^\mu_\nu \frac{dx^\nu}{d\tau} \tag{117}$$

Various derivatives of the electromagnetic field strength tensor F can yield useful information. For instance, two of Maxwell's equations, namely (99) and (100), can be attained by (loosely speaking) taking the *four-curl* of F:

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0 \tag{118}$$

Problem 46: Prove that $\vec{\nabla} \cdot \vec{B} = 0$ by choosing the indices $\alpha = 1$, $\beta = 2$, $\gamma = 3$ in equation (118). Do any other choices of α, β, γ give the same result?

Problem 47: Prove that the Maxwell equation $\vec{\nabla} \times \vec{E} = -\partial \vec{B}/\partial t$ is contained in (118).

Problem 48: Express F as a type $(2,0)$ tensor. Answer:

$$F^{\mu\nu} = \begin{bmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{bmatrix} \tag{119}$$

Problem 49: Express F as a type $(1,1)$ tensor. Answer:

$$F^\mu_\sigma = \begin{bmatrix} 0 & E_x/c & E_y/c & E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{bmatrix} \tag{120}$$

Problem 50: Consider two inertial reference frames in relative motion at speed v in the x-direction. Use the transformation $F'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}$ to determine the components of F' in terms of the unprimed fields. (Note this is not the *whole* story, because the components of x^μ must also be Lorentz-transformed.) Answer:

$$E'_x = E_x \qquad B'_x = B_x$$

$$E'_y = \gamma\left(E_y - vB_z\right) \qquad B'_y = \gamma\left(B_y + vE_z/c^2\right)$$
$$E'_z = \gamma\left(E_z + vB_y\right) \qquad B'_z = \gamma\left(B_z - vE_y/c^2\right)$$

Problem 51: For a boost in an arbitrary direction with velocity $\vec{v}$, convince yourself that the Lorentz transformation $F'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}$ generalizes to:

$$E'_\parallel = E_\parallel \qquad B'_\parallel = B_\parallel$$

$$\vec{E}'_\perp = \gamma\left(\vec{E}_\perp - \vec{v}\times\vec{B}\right) \qquad \vec{B}'_\perp = \gamma\left(\vec{B}_\perp + \frac{\vec{v}\times\vec{E}}{c^2}\right)$$

Problem 52: Prove that the tensor contraction, or *inner product* of the electromagnetic field tensor is equatl to the Lorentz invariant $2(B^2 - E^2/c^2)$. In other words, prove

$$F_{\mu\nu}F^{\mu\nu} = 2\left(B^2 - \frac{E^2}{c^2}\right) . \tag{121}$$

Problem 53: Another Lorentz invariant buried in the electromagnetic field strength tensor is the quantity $-4\vec{E}\cdot\vec{B}/c$, called a *pseudoscalar* invariant. Prove:

$$\frac{1}{2}\epsilon_{\alpha\beta\gamma\rho}F^{\alpha\beta}F^{\gamma\rho} = -\frac{4}{c}\vec{E}\cdot\vec{B} \tag{122}$$

Problem 54: Yet another Lorentz invariant is contained in the determinant of F. Prove that

$$\det\left(F\right) = \frac{c}{c^2}\left(\vec{E}\cdot\vec{B}\right)^2 . \tag{123}$$

Problem 55: The Lorentz force law $\vec{F} = q\vec{v}\times\vec{B}$ is only correct in the nonrelativistic limit. Use equation (117) to derive the true force laws for a boost along x. Answer:

$$\frac{d\left(E - E_0\right)}{dt} = \frac{dT}{dt} = \frac{q\vec{E}\cdot\vec{v}}{\gamma^2} \qquad m\left(\frac{d^2\vec{x}}{d\tau^2}\right) = q\gamma\left(\vec{E} + \vec{v}\times\vec{B}\right)$$

Problem 56: Solve the relativistic equation of motion (117) for a charged particle initially at rest and immersed in the constant electric field $\vec{E} = E\hat{x}$. Obtain solutions in terms of both t and τ. (You are free to check that they are equivalent.) Answer:

$$v\left(t\right) = c\frac{qEt}{\sqrt{m^2c^2 + q^2E^2t^2}} \qquad x\left(t\right) = \frac{c}{qE_0}\sqrt{m^2c^2 + q^2E^2t^2} - \frac{mc^2}{qE_0} + x_0$$

$$\frac{d^2U^\sigma}{d\tau^2} = \left(\frac{qE}{mc}\right)^2 U^\sigma \qquad U^t = \cosh\left(\frac{qE}{mc}\tau\right) \qquad U^x = \sinh\left(\frac{qE}{mc}\tau\right)$$

$$t\left(\tau\right) = \frac{mc}{qE}\sinh\left(\frac{qE}{mc}\tau\right) \qquad x\left(\tau\right) = \frac{mc^2}{qE}\cosh\left(\frac{qE}{mc}\tau\right) - \frac{mc^2}{qE} + x_0$$

Problem 57: Solve the relativistic equation of motion (117) for a charged particle with arbitrary initial conditions and immersed in the magnetic field $\vec{B} = B\hat{x}$. Answer:

$$U^x\left(\tau\right) = \dot{x}_0 \qquad U^y\left(\tau\right) = A\sin\left(\frac{qB}{m}\tau + \phi\right) \qquad U^z\left(\tau\right) = A\cos\left(\frac{qB}{m}\tau + \phi\right)$$

$$U^0 = c\frac{dt}{d\tau} \qquad t = \tau\sqrt{1 + \frac{\dot{x}_0^2 + A^2}{c^2}} \qquad x = x_0 + \dot{x}_0\tau$$

$$y = y_0 - \frac{Am}{qB}\cos\left(\frac{qB}{m}\tau + \phi\right) \qquad z = z_0 + \frac{Am}{qB}\sin\left(\frac{qB}{m}\tau + \phi\right)$$

Problem 58: A point charge of magnitude q stays fixed at the origin of a reference frame x^μ. Write down the the electric and magnetic fields $\vec{E}$, $\vec{B}$ for all points in x^μ. A second reference frame moving along the x direction at speed v observes the same point charge. Use the Lorentz transformation to find the electric and magnetic fields $\vec{E}'$, $\vec{B}'$ at all points in the boosted frame $(x')^\mu$. Answer:

$$\vec{E} = \frac{q}{4\pi\epsilon_0}\frac{\vec{x}}{|\vec{x}|^3} \qquad \vec{B} = 0$$

$$\vec{E}' = \frac{q}{4\pi\epsilon_0}\frac{\left(\gamma\left(x'-vt'\right), \gamma y', \gamma z'\right)}{\left(\gamma^2\left(x'-vt'\right)^2 + y'^2 + z'^2\right)^{3/2}}$$

$$\vec{B}' = \frac{v}{c^2}\frac{q}{4\pi\epsilon_0}\frac{\left(0, \gamma z', -\gamma y'\right)}{\left(\gamma^2\left(x'-vt'\right)^2 + y'^2 + z'^2\right)^{3/2}}$$

Problem 59: The same point charge of magnitude q stays fixed at the origin of a reference frame x^μ. A second reference frame moves along the x-axis at ultra-relativistic speed where $v \to c$. Show that the observed fields $\vec{E}'$ and $\vec{B}'$ are:

$$\vec{E}' = \delta\left(x' - ct'\right)\frac{2q}{4\pi\epsilon_0}\frac{\left(0, y', z'\right)}{y'^2 + z'^2}$$

$$\vec{B}' = \delta\left(x' - ct'\right)\frac{v}{c^2}\frac{2q}{4\pi\epsilon_0}\frac{\left(0, z', -y'\right)}{y'^2 + z'^2}$$

Note that the above fields are confined to a two-dimensional plane that extends orthogonally to the direction of motion, called an electromagnetic *shock wave*. Hint:

$$\lim_{\beta\to 1}\frac{1}{\sqrt{1-\beta^2}} f\left(\frac{x-ct}{1-\beta^2}\right) = \delta\left(x-ct\right)\int_{-\infty}^{\infty} f\left(w\right) dw \tag{124}$$

Problem 60: Derive equation (124) using the theory of distributions.

6.3 Four-Current and Gauss-Ampere Law

The core of electromagnetism need not be recast to accomodate special relativity, however the notion of warped spacetime must be integrated into theory. For instance, the unit charge is invariant for all reference frames, but the charge density and current density clearly depend on the frame of reference. We therefore construct a four-vector called the *four-current* as:

$$J^\nu = \left(\rho c, \rho\vec{v}\right) \tag{125}$$

Note ρ is the electric charge density, and the quantity $\rho\vec{v}$ is defined as the current density $\vec{J}$. The nonrelativistic continuity equation (102) may be recovered by taking the four-gradient of the four-current:

$$\partial_\mu J^\mu = \frac{\partial J^\alpha}{\partial x^\alpha} = \frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot\vec{J} = 0 \tag{126}$$

Problem 61: Determine the components of the four-current in a primed reference frame that is Lorentz-boosted along the x-direction at speed u. Answer:

$$J'_x = \gamma\left(J_x - u\rho\right) \qquad\qquad \rho' = \gamma\left(\rho - uJ_x/c^2\right)$$

Problem 62: The quantity $J^\mu J_\mu$ must be Lorentz-invariant. Use this fact to derive the relation

$$\rho = \gamma\rho_0 \; ,$$

where ρ_0 is the invariant *proper density*.

Consider a space that has electric and magnetic fields $\vec{E}$ and $\vec{B}$ defined everywhere with some arbitrary arrangement of electric charge densities ρ and currents $\vec{J}$. We'll show using Hamilton's principle that classical electromagnetism can be derived by minimizing the action:

$$S = \int \left[-\frac{1}{4\mu_0}F_{\mu\nu}F^{\mu\nu} - J^\mu A_\mu\right] d^4x \tag{127}$$

The quantity in square brackets is the Lagrangian density, denoted $\mathcal{L}$, with corresponding Euler-Lagrange equation for fields (derived previously):

$$\frac{\partial}{\partial x_\nu}\left(\frac{\partial\mathcal{L}}{\partial\left(\partial A_\mu/\partial x_\nu\right)}\right) - \frac{\partial\mathcal{L}}{\partial A_\mu} = 0 \tag{128}$$

Simplifying (127) and picking out $\mathcal{L}$, we have

$$\mathcal{L} = -\frac{1}{2\mu_0}\left(\partial_\mu A_\nu\partial^\mu A^\nu - \partial_\nu A_\mu\partial^\mu A^\nu\right) - J^\mu A_\mu \; ,$$

where (114) has replaced explicit mention of F. Ultimately we must crank $\mathcal{L}$ through equation (128); it helps to first solve the mini-problem:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\left(\partial A_\mu/\partial x_\nu\right)} &= \frac{\partial}{\partial\left(\partial A_\mu/\partial x_\nu\right)}\left(-\frac{1}{2\mu_0}\left(\partial_\sigma A_\lambda\partial^\sigma A^\lambda - \partial_\lambda A_\sigma\partial^\sigma A^\lambda\right)\right) \\ &= -\frac{1}{2\mu_0}2\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) \\ &= -\frac{1}{\mu_0}F^{\mu\nu} = \frac{1}{\mu_0}F^{\nu\mu}\end{aligned} \tag{129}$$

Note that the flat metric η is used to lower all indices before the derivative is evaluated.

The remaining part of (128) is trivial to evaluate, resulting in the famous *Gauss-Ampere law*:

$$\partial_\nu F^{\nu\mu} = -\mu_0 J^\mu \tag{130}$$

The Gauss-Ampere law isn't a new law, but in fact contains the *other* two Maxwell's equations, namely (98) and (101), that weren't covered by (118). The left side of (130) is a four-vector, which after simplifying becomes:

$$\partial_\nu F^{\nu\mu} = \begin{bmatrix} -\vec{\nabla}\cdot\vec{E}/c \\ c^{-2}\partial\vec{E}/\partial t - \vec{\nabla}\times\vec{B}\end{bmatrix} = \begin{bmatrix} -\mu_0 c\rho \\ -\mu_0\vec{J}\end{bmatrix} = -\mu_0 J^\mu \tag{131}$$

Problem 63: Prove equation (131), which completes our repacking of Maxwell's equations from four down to two.

6.4 Transformations and Gauges

Classical electromagnetism is nicely contained in various derivatives of the electromagnetic field strength tensor F, which itself is derived from the four-potential A. It's natural to wonder what may be the restrictions on the four-potential.

Lorenz Gauge

The continuity equation $\partial_\mu J^\mu = 0$ automatically implies that the four-potential may be translated by a harmonic scalar function f^μ that obeys $\partial^\mu \partial_\mu f = 0$ as

$$A^\mu \to A^\mu + f^\mu \, , \tag{132}$$

with no physical consequences. The translation (132) may also be written

$$\partial_\mu A^\mu = 0 \, , \tag{133}$$

which is known as the *Lorenz gauge-fixing condition.* (Note 'Lorenz' is not a misspelling of 'Lorentz' - these are two different names.)

Problem 64: Using the definition $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, prove that the electromagnetic field strength tensor F is invariant under the translation (132).

Problem 65: Show that, in ordinary vector notation, the Lorenz gauge-fixing condition is:

$$\frac{1}{c^2}\frac{\partial \phi}{\partial t} + \vec{\nabla} \cdot \vec{A} = 0$$

Problem 66: In the Lorenz gauge, derive the inhomogeneous wave equations:

$$-\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} + \nabla^2 \phi = -\frac{\rho}{\epsilon_0} \tag{134}$$

$$-\frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} + \nabla^2 \vec{A} = -\mu_0 \vec{J} \tag{135}$$

Problem 67: Check that the d'Alembertian operator

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$$

allows the inhomogeneous wave equations (134) and (135) to be written in terms of the four-current as

$$\Box A^\alpha = \mu_0 J^\alpha \, . \tag{136}$$

Coulomb Gauge

Another way to restrict the vector potential $\vec{A}$ is to use the *Coulomb* (a.k.a. radiation or transverse) *gauge fixing condition*, which is simply

$$\vec{\nabla} \cdot \vec{A} = 0 \, .$$

The equations of motion (110) and (111) become:

$$\nabla^2 \phi = -\frac{\rho}{\epsilon_0} \tag{137}$$

$$\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = -\mu_0 \vec{J} + \vec{\nabla}\left(\frac{1}{c^2}\frac{\partial \phi}{\partial t}\right) \tag{138}$$

It can be shown that the final term in (138) can be restated as μ_0 times the longitudinal component of the current density J_l such that $\vec{\nabla} \times \vec{J_l} = 0$. The transverse component of the current density J_t has the property $\vec{\nabla} \cdot \vec{J_t} = 0$ and $\vec{J} = \vec{J_l} + \vec{J_t}$ according to Helmholtz's theorem. It follows that (138) may be restated as:

$$\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = -\mu_0 \vec{J_t}$$

6.5 Electromagnetic Stress-Energy Tensor

The electromagnetic field strength tensor $F^{\mu\nu}$ encapsules all information about the electric and magnetic fields in spacetime. Another useful quantity is the stress-energy tensor for electromagnetism, which contains information on various products of $\vec{E}$ and $\vec{B}$: the Poynting vector, the energy density, and so on. In flat space, the Lagrangian density is

$$\mathcal{L} = -\frac{1}{4\mu_0} F_{\mu\nu}F^{\mu\nu} - A_\mu J^\mu \,.$$

The stress-energy tensor shall be calculated from equation (97), namely

$$T^{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial\left(\partial_\mu A_\lambda\right)}\partial^\nu A_\lambda + \eta^{\mu\nu}\mathcal{L} + \partial_\lambda \mathcal{K}^{[\lambda\mu]\nu} \,.$$

Subsituting $\mathcal{L}$ into the equation for $T^{\mu\nu}$, we get

$$T^{\mu\nu} = -\frac{1}{\mu_0}F^{\mu\lambda}\partial^\nu A_\lambda - \frac{1}{4\mu_0}\eta^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} - \eta^{\mu\nu}A_\lambda J^\lambda + \partial_\lambda \mathcal{K}^{[\lambda\mu]\nu} \,,$$

which would have no chance of being symmetric or traceless without the added $\mathcal{K}$ tensor. Moreover, $\mathcal{K}$ ensures gauge-invariance in $T^{\mu\nu}$, which we'd expect because $F^{\mu\nu}$ is gauge-invariant. It follows that the choice for $\mathcal{K}$ obeys

$$\partial_\lambda \mathcal{K}^{[\lambda\mu]\nu} = \frac{1}{\mu_0}F^{\mu\lambda}\partial_\lambda A^\nu \,.$$

The electromagnetic stress-energy tensor in flat space is therefore

$$T^{\mu\nu}_{EM} = \frac{1}{\mu_0}\left[F^{\mu\lambda}F^{\nu}_{\lambda} - \frac{1}{4}\eta^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right] + \frac{1}{\mu_0}\eta^{\mu\nu}A_\lambda J^\lambda \,. \tag{139}$$

Problem 68: Check that the homogeneous ($J^\lambda = 0$) electromagnetic stress-energy tensor $T^{\mu\nu}_{EM}$ is invariant under gauge transformations $A_\mu \to A_\mu + \partial_\mu \psi$.

Problem 69: Check that the homogeneous ($J^\lambda = 0$) electromagnetic stress-energy tensor $T^{\mu\nu}_{EM}$ has zero divergence, or

$$\partial_\mu T^{\mu\nu}_{EM} = 0 \,. \tag{140}$$

Components of the Electromagnetic Stress-Energy Tensor

Determining the components of $T^{\mu\nu}_{EM}$ is a matter of brute-force calculation based on our previous achievements. In space not enclosing charges or currents, the homogeneous version ($J^\lambda = 0$) reads

$$T^{\mu\nu} = \begin{bmatrix} \frac{1}{2}\left(\epsilon_0 E^2 + \frac{1}{\mu_0}B^2\right) & S_x/c & S_y/c & S_z/c \\ S_x/c & -\sigma_{xx} & -\sigma_{xy} & -\sigma_{xz} \\ S_y/c & -\sigma_{yx} & -\sigma_{yy} & -\sigma_{yz} \\ S_z/c & -\sigma_{zx} & -\sigma_{zy} & -\sigma_{zz} \end{bmatrix} , \tag{141}$$

where S_j are components of the *Poynting vector*

$$\vec{S} = \frac{1}{\mu_0}\vec{E} \times \vec{B} , \tag{142}$$

and the *Maxwell stress tensor* σ_{ij} is

$$\sigma_{ij} = \epsilon_0 E_i E_j + \frac{1}{\mu_0} B_i B_j - \frac{1}{2}\left(\epsilon_0 E^2 + \frac{1}{\mu_0}B^2\right)\delta_{ij} . \tag{143}$$

Problem 70: Check that the homogeneous ($J^\lambda = 0$) electromagnetic stress-energy tensor $T^{\mu\nu}_{EM}$ is traceless.

7 Fluids in Flat Space

7.1 Number-Flux and Conservation of Matter

In the viewpoint of special and general relativity, matter is modeled as differentiable media - a *fluid* in spacetime. This model doesn't comply with quantum mechanics below certain length scales, which is for all practical purposes is irrelevant, as we shall concern over large length scales. (Those who *are* concerned with this might be looking for a quantum theory of gravity, or better yet, a *Grand Unified Theory*, which itself may or may not turn out to be practical.)

Any element of a fluid can be described by two quantities: it's four-velocity $U^\mu(x^\alpha)$, and its number density $n(x^\alpha)$. Note that U^μ is a four-vector, and n is a scalar function; the product of which is defined as the *number-flux four-vector*:

$$N^\mu\left(x^\alpha\right) = n\left(x^\alpha\right)U^\mu\left(x^\alpha\right) \tag{144}$$

One subtlety in this model is the unit volume is not relativistically invariant, although the number of particles per unit volume is invariant. It makes sense to define the 'proper' unit volume in the rest frame.

For nonrelativistic speeds ($\gamma \approx 1$), the flux four-vector becomes

$$N^\mu = \left(c \cdot n\left(\vec{x}, t\right), \vec{J}\left(\vec{x}, t\right)\right) ,$$

where $\vec{J}(\vec{x}, t)$ is the nonrelativistic number-flux three-vector. By the same considerations that give rise to the electrodynamic continuity equation (102), and we arrive at

$$\frac{\partial n}{\partial t} + \vec{\nabla} \cdot \vec{J} = 0 \qquad \rightarrow \qquad \partial_\mu N^\mu = 0 ,$$

which assures local conservation of matter.

7.2 Stress-Energy Tensor for Matter

The most general definition of the stress-energy tensor reads: '$T^{\mu\nu}$ is the flux of the four-momentum $P^\mu = (E/c, \vec{p})$ across a surface of constant $x^\nu = (ct, \vec{x})$.' Following this definition, the components of $T^{\mu\nu}$ are:

$$T^{00} = \frac{\Delta E/c}{d^3x} = c^{-2} \cdot \text{Energy Density} \tag{145}$$

$$T^{i0} = \frac{\Delta P^i/c}{d^3x} = \text{Momentum Density} \tag{146}$$

$$T^{0i} = \frac{\Delta E/c}{\Delta t} \frac{1}{\Delta x^j \Delta x^k} = \text{Energy Flux} \tag{147}$$

$$T^{ij} = \frac{\Delta P^i/c}{\Delta t} \frac{1}{\Delta x^i \Delta x^k} = F \perp j = \text{Shear Stress} \tag{148}$$

$$T^{ii} = \frac{\Delta P^i/c}{\Delta t} \frac{1}{\Delta x^j \Delta x^k} = \text{Pressure} \tag{149}$$

Problem 71: Check that the stress-energy tensor for matter obeys $T^{0i} = T^{i0}$ by showing that (146) and (147) are equivalent.

7.3 Isolated Particle

A single particle of mass m that traces out the trajectory $\vec{x}_p$ with velocity components v^μ has

$$T^{\mu\nu} = m \frac{v^\mu v^\nu}{\sqrt{1 - v^2/c^2}} \delta(\vec{x} - \vec{x}_p) \ .$$

Since we know $E = \gamma mc^2$, we may instead write

$$T^{\mu\nu} = E \frac{v^\mu v^\nu}{c^2} \delta(\vec{x} - \vec{x}_p) \ , \tag{150}$$

where E may also be replaced via $E^2 = p^2c^2 + m^2c^4$.

Problem 72: What is $T^{\mu\nu}$ for a particle at rest? Answer:

$$T^{\mu\nu} = mc^2 \delta(\vec{x}_p) \delta^{00}$$

7.4 Perfect Fluid

Consider a fluid that is in thermodynamic equilibrium, also known as a *perfect fluid*. Any element of the fluid has a four-velocity U^μ, and the stress-energy tensor takes the form

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) U^\mu U^\nu + p\eta^{\mu\nu} \ , \tag{151}$$

where ρ is the mass-density of the fluid and p is the hydrostatic pressure. In the rest frame of the fluid, the four-velocity is simply $U^\mu = (c, 0)$, and the stress-energy tensor reduces to

$$T^{\mu\nu} = \text{diag}\{c^2\rho, p, p, p\} \ . \tag{152}$$

Problem 73: A perfect fluid in its rest frame is mapped by coordinates x^μ. A Lorentz-boosted frame x'^μ moves with speed v along the x-axis of the unprimed (rest) frame. Find the components of $(T')^{\mu\nu}$. Answer:

$$(T')^{00} = \frac{c^2\rho + \beta^2 p}{1-\beta^2} \qquad\qquad (T')^{11} = \frac{v^2\rho + p}{1-\beta^2}$$

Problem 74: Verify the property $\partial_\mu T^{\mu\nu} = 0$ for the perfect fluid.

8 Uniform Acceleration

8.1 Uniformly-Accelerated Particle

Consider an inertial reference frame in where someone observes a uniformly-accelerated particle moving along the x-axis. Our method for analyzing this scenario, which shall be employed for more complicated situations, is to treat the accelerated reference frame as an *instantaneous* inertial frame. That is, employ the full apparatus of special relativity with the understanding that the relative frame speed v is no longer constant, but evolves according to $dv/dt = \tilde{a}$. Since the y- and z- directions don't have consequence for this analysys, we'll primarily work with two-vectors instead of four-vectors.

Cover the inertial frame with coordinates (a.k.a. place observers at) x^μ and the accelerated frame with coordinates $(x')^\mu$. At any instant, the coordinate systems are related by the Lorentz transformation

$$(x')^\mu = \Lambda^\mu_\nu x^\nu \; .$$

Meanwhile, the two-velocity of the particle is given by equation (37) as

$$U^\mu = \left(c\frac{dt}{d\tau}, \frac{dx}{d\tau}\right) = (\gamma c, \gamma v) \; ,$$

where γ and v are both functions of τ, the proper time of the accelerated reference frame. If the particle could wear a wristwatch, it would keep track of τ.

For a boost along the x-direction, the nontrivial components of Λ are

$$\Lambda^0_0 = \Lambda^1_1 = \cosh(\phi(\tau)) \qquad\qquad \Lambda^1_0 = \Lambda^0_1 = -\sinh(\phi(\tau)) \; ,$$

where ϕ is the rapidity corresponding to speed v. Equivalently, this is

$$\Lambda^0_0 = \Lambda^1_1 = \gamma(\tau) \qquad\qquad \Lambda^1_0 = \Lambda^0_1 = -\gamma(\tau)\frac{v(\tau)}{c} \; ,$$

where we identify the velocity as $v = c\tanh(\phi(\tau))$. To keep going, notice the two-acceleration is a τ-derivative of the two-velocity, which for the x-accelerating particle reads

$$a^\mu = c\frac{d}{d\tau}(\cosh\phi, \sinh\phi) = c\left(\sinh\phi\frac{\partial\phi}{\partial\tau}, \cosh\phi\frac{\partial\phi}{\partial\tau}\right) \; .$$

The acceleration must be constant by construction, which has consequence for $d\phi/d\tau$. Calculate the Lorentz-transformation of a^μ by using

$$(a')^\mu = \Lambda^\mu_\nu a^\nu \; ,$$

and focus on the $(a')^x$ component. In particular, we have

$$\tilde{a} = \left(a'\right)^x = \Lambda^x_\nu a^\nu = \Lambda^1_0 a^0 + \Lambda^1_1 a^1 = c\frac{\partial \phi}{\partial \tau} \, ,$$

so evidently

$$\phi\left(\tau\right) = \tilde{a}\tau/c + \phi_0 \, ,$$

where we set ϕ_0 equal to zero by convention.

The velocity two-vector of the uniformly accelerated particle along x is therefore

$$U^\mu = c\left(\cosh\left(\tilde{a}\tau/c\right), \sinh\left(\tilde{a}\tau/c\right)\right) \, , \tag{153}$$

and after integrating with respect to τ we have the position two-vector of the particle:

$$x^0 = ct\left(\tau\right) = \frac{c^2}{\tilde{a}}\sinh\left(\tilde{a}\tau/c\right) \tag{154}$$

$$x^1 = x\left(\tau\right) = \frac{c^2}{\tilde{a}}\cosh\left(\tilde{a}\tau/c\right) - \frac{c^2}{\tilde{a}} \tag{155}$$

Note the integration constants are chosen such that both reference frames coincide such that $x = x' = 0$ at $t = \tau = 0$.

Problem 75: The equations of motion for a uniformly-accelerated particle by a force F were first found when we covered relativistic mechanics, and by equation (21) we found

$$v\left(t\right) = \frac{c}{\sqrt{1 + m^2c^2/F^2t^2}} \qquad\qquad x\left(t\right) = \frac{c}{F}\sqrt{F^2t^2 + m^2c^2} - \frac{mc^2}{F}$$

Using $F = m\tilde{a}$, manipulate equations (153), (154), and (155) to reproduce the relations (21) rewritten above. Hint: $\cosh\left(\sinh^{-1}\left(x\right)\right) = \sqrt{1 + x^2}$

Problem 76: In the nonrelativistic limit, show that the equations of motion (154), (155) for the uniformly-accelerated particle reduce to the usual formulae delivered by Newton's laws. Answer: For $(\tilde{a}\tau/c)$ small,

$$t\left(\tau\right) = \tau \qquad\qquad x\left(\tau\right) = \frac{c^2}{\tilde{a}}\left(1 + \frac{1}{2}\left(\tilde{a}\tau/c\right)^2 + ... - 1\right) \approx \frac{1}{2}\tilde{a}\tau^2 \, .$$

8.2 Uniformly-Accelerated Rocket

Now extend the analysis to handle uniformly-accelerated objects of finite size. In this section we shall replace the previous example's particle with a linear rocket that is uniformly accelerated along (and aligned with) the x-direction.

To calibrate the stationary and accelerated reference fames, suppose the floor of the rocket is located at $x = 0$ at $t = \tau = 0$. As time evolves, the floor occupies the position $x(t)$ according to a stationary observer. In the rocket's frame, the floor remains at $x' = 0$ for all τ, the proper time 'experienced' at the floor. Denote ξ as any given height above the floor as observed from inside the rocket. Our goal is to find new equations of motion (according to a stationary observer) that apply to the whole interior of the rocket.

The key insight is two notice that there are *two* contributions to the total boost at point $x' = \xi$. First, the floor of the rocket has equations obeying (154) and (155). Second, the point ξ has its own Lorentz boost relative to the floor at $x' = 0$. The components of $x^\mu(\xi)$, in accordance with equations (5) and (6), are

$$t\,(\tau,\xi) = \gamma\left(t_0 + v x_0/c^2\right) + t_{\text{floor}}$$

$$x\,(\tau,\xi) = \gamma\,(x_0 + v t_0) + x_{\text{floor}}\ ,$$

and by construction we have

$$\gamma = \cosh(\tilde{a}\tau/c) \qquad\qquad t_0 = 0 \qquad\qquad x_0 = \xi\ .$$

The equations of motion for all points within the rocket are the components of $x^\mu(\xi)$, namely

$$t\,(\tau,\xi) = \frac{1}{c}\left(\frac{c^2}{\tilde{a}} + \xi\right)\sinh\left(\tilde{a}\tau/c\right) \tag{156}$$

$$x\,(\tau,\xi) = \left(\frac{c^2}{\tilde{a}} + \xi\right)\cosh\left(\tilde{a}\tau/c\right) - \frac{c^2}{\tilde{a}}\ . \tag{157}$$

Problem 77: Solve the equations of motion (156), (157) for $\tau(x,t)$ and $\xi(x,t)$. Answer:

$$\tau\,(x,t) = \frac{c}{\tilde{a}}\text{arctanh}\left(\frac{ct}{x + c^2/\tilde{a}}\right)$$

$$\xi\,(x,t) = -\frac{c^2}{\tilde{a}} + \sqrt{\left(x + c^2/\tilde{a}\right)^2 - (ct)^2}$$

8.3 Rindler Metric

The equations of motion (156) and (157) bring out some of the exotic predictions made by special relativity. Begin by computing the spacetime interval for the space embedded in the accelerating rocket: the general procedure is to determine each $(dx^\mu)^2$ in terms of $(dx'^\mu)^2$, and then assemble the sum $dS^2 = -(dx^0)^2 + (dx^1)^2$. The factors that remain attached to the terms $dx^\mu dx^\nu$ become the metric components $g^{\mu\nu}$. Explicitly, we start with

$$dx^0 = d\tau\left(c + \frac{\xi\tilde{a}}{c}\right)\cosh\left(\tilde{a}\tau/c\right) + d\xi\sinh\left(\tilde{a}\tau/c\right)$$

$$dx^1 = d\tau\left(c + \frac{\xi\tilde{a}}{c}\right)\sinh\left(\tilde{a}\tau/c\right) + d\xi\cosh\left(\tilde{a}\tau/c\right)\ ,$$

and we find

$$dS^2 = -c^2 d\tau^2\left(1 + \frac{\xi\tilde{a}}{c^2}\right)^2 + d\xi^2\ . \tag{158}$$

Equation (158) is called the *Rindler spacetime interval*, or equivalently the Rindler metric. We'll pick out three surprising features:

- The local clock speed varies with the coordinate ξ.
- The local acceleration varies with the coordinate ξ.
- Rindler spacetime does not cover all of Minkowski spacetime.

Problem 78: From equation (158), write out the components of the Rindler metric $g_{\mu\nu}$ in block form. Answer:

$$g_{\mu\nu} = \begin{bmatrix} -\left(1 + \xi\tilde{a}/c^2\right)^2 & 0 \\ 0 & 1 \end{bmatrix} \tag{159}$$

Inverse Rindler Metric

It's a useful exercise to determine the Rindler metric as a type $(2,0)$ tensor. Using the definition (47) we begin with

$$g^{\mu\alpha} g_{\alpha\nu} = \delta^{\mu}_{\nu} \, .$$

Notice that setting $\mu = \nu$ involves a sum over the total number of dimensions in the problem, which for Rindler space is $1+1 = [\tau] + [\xi] = 2$. Inserting (159) into the above, arrive at the statement

$$-(1+\xi\tilde{a}/c^2)^2 g^{00} + g^{11} = 2 \, ,$$

and conclude

$$g^{\mu\nu} = \begin{bmatrix} -1/\left(1+\xi\tilde{a}/c^2\right)^2 & 0 \\ 0 & 1 \end{bmatrix} . \tag{160}$$

8.4 Properties of Rindler Space

Local Clock Speed

The uniformly-accelerated frame is indeed one single reference frame, yet the spacetime interval (158) tells us the rate of time flow varies per spatial coordinate ξ. Recall the floor of the rocket is at $x' = 0$, and the proper time at the floor is always τ. At heights $x' = \xi$ above the floor, let the corrpsonding time variable be τ_ξ.

By equation (158), we have

$$d\tau_\xi = d\tau \left(1 + \frac{\xi\tilde{a}}{c^2}\right) \, , \tag{161}$$

telling us that $d\tau$ is smaller than any other $d\tau_\xi$. That is, if you are standing up in an accelerating rocket, your head is aging faster than your feet. This effect is called the *gravitational redshift*.

Photon Redshift

An accelerating light source will emit photons with varying wavelengths according to a stationary observer. In the case the $v \ll c$, the relativisic doppler equation (33) becomes

$$f = f' \left(1 + \frac{v}{c}\right) \, ,$$

where f is the observed frequency of a photon. The sign on v determines whether the photon is redshifted or blueshifted. For the case of uniform acceleration, v may be replaced with $-\xi\tilde{a}/c$, recovering equation (161) (remembering that $f = 1/T$).

Local Acceleration

The acceleration two-vector of a point $x' = \xi$ above the floor of the rocket is defined as $a^\mu_\xi = d^2 x^\mu / d\tau_\xi^2$. (The floor itself accelerates at a constant $\tilde{a}$.) Using the chain rule and equation (161), arrive at the relations

$$a^0_\xi = \left(\frac{\tilde{a}}{1+\xi\tilde{a}/c^2}\right) \sinh\left(\xi\tilde{a}/c\right) \qquad a^1_\xi = \left(\frac{\tilde{a}}{1+\xi\tilde{a}/c^2}\right) \cosh\left(\xi\tilde{a}/c\right) \, ,$$

where the second derivatives of $t(\tau)$ and $x(\tau)$ have been derived from the equations of motion (156) and (157), respectively.

The norm (Minkowski dot product onto self) of a^μ gives the acceleration of an observer at the fixed coordinate $x' = \xi$. Denoting this acceleration $\tilde{a}_\xi$, we have

$$(\tilde{a}_\xi)^2 = -\left(a_\xi^t\right)^2 + \left(a_\xi^x\right)^2 = \left(\frac{\tilde{a}}{1+\xi\tilde{a}/c^2}\right)^2 ,$$

implying

$$\tilde{a}_\xi = \frac{\tilde{a}}{1+\xi\tilde{a}/c^2} . \tag{162}$$

Problem 79: Recover equation (162) by Taylor-expanding the equations of motion (156), (157) and picking out the effective acceleration. Answer:

$$\tau \approx \frac{t}{(1+\xi\tilde{a}/c^2)} \qquad\qquad x\,(t) \approx \xi + \frac{t^2}{2}\left(\frac{\tilde{a}}{1+\xi\tilde{a}/c^2}\right)$$

Wordline of Stationary Object

Problem 80: A stationary object is located in the non-accelerated reference frame at $x^\mu(x_0)$, where $x_0 > -c^2/\tilde{a}$. According to the accelerated coordinate system, the object is moving along a worldline $\xi(\tau)$. Determine $\xi(\tau)$ from the equations of motion (156) and (157), and show that the solution is asymptotic to $\xi = -c^2/\tilde{a}$ for very large τ. Answer:

$$\xi\,(\tau) = \frac{x_0 + c^2/\tilde{a}}{\cosh\left(\tilde{a}\tau/c\right)} - \frac{c^2}{\tilde{a}} \qquad\qquad \xi\,(\tau\to\pm\infty) = -\frac{c^2}{\tilde{a}}$$

Problem 81: The same stationary object is located in the non-accelerated reference frame at $x^\mu(x_0)$, where $x_0 > -c^2/\tilde{a}$. The two-velocity of the object is defined as the derivative of the position two-vector with respect to the object's proper time, which in this case is simply t:

$$\left(V^\tau, V^\xi\right) = \left(c\frac{d\tau}{dt}, \frac{d\xi}{dt}\right)$$

In Rindler space, the norm of a two-vector is given by $||U\cdot U|| = g^{\mu\nu}U_\mu U_\nu$, where $g^{\mu\nu}$ is the Rindler metric. Compute the components of V^μ in terms of τ, ξ, and then show that $V^\mu V_\mu = -c^2$. Answer:

$$V^\mu = c\left(\frac{\cosh\left(\tilde{a}\tau/c\right)}{1+\xi\tilde{a}/c^2}, -\sinh\left(\tilde{a}\tau/c\right)\right)$$

Many Observers with Constant Acceleration

Problem 82: Let us now find a set of observers that experience one and the same acceleration $\tilde{a}$. At time $t=0$, each observer is defined to be in some rest position $\bar{x}$, and as time evolves, each carries its own proper time $\theta(x,t)$. Determine the equations of the worldline $t(\theta,\bar{x})$, $x(\theta,\bar{x})$ of any such observer. Calculate the spacetime interval dS^2 in terms of the coordinates θ, $\bar{x}$. Answer:

$$t = \frac{c}{\tilde{a}}\sinh\left(\tilde{a}\theta/c\right)$$

$$x = \bar{x} + \frac{c^2}{\tilde{a}}\left(\cosh\left(\tilde{a}\theta/c\right) - 1\right)$$

$$dS^2 = -c^2dt^2 + dx^2 = -c^2d\theta^2 + 2c\sinh\left(\tilde{a}\theta/c\right)d\bar{x}d\theta + d\bar{x}^2$$

Uniform Electric Field

Problem 83: In 1+1 dimensions, the only nonvanishing component of the electromagnetic field strength tensor is, for instance $F_{tx} = -F_{xt} = E/c = constant$. (The magnetic field does not exist when there is only one spatial dimension.) Determine the components of $F_{\tau\xi}$, the electromagnetic field strength tensor as observed in an accelerated frame obeying the equations of motion (156), (157). Note the result E' should be non-uniform. Answer:

$$F_{\tau\xi} = \frac{\partial x^\mu}{\partial \tau}\frac{\partial x^\nu}{\partial \xi}F_{\mu\nu} \qquad\qquad E' = E\left(1 + \frac{\xi\tilde{a}}{c^2}\right)$$

Alternative Rindler Metric

Problem 84: Consider the equations of motion

$$t = \frac{c}{\tilde{a}}e^{\xi\tilde{a}/c^2}\sinh\left(\tilde{a}\lambda/c\right) \qquad\qquad x = \frac{c^2}{\tilde{a}}e^{\xi\tilde{a}/c^2}\cosh\left(\tilde{a}\lambda/c\right)\ . \tag{163}$$

Determine the alternative Rindler metric by evaluating the Minkowski spacetime interval $dS^2 = \eta_{\mu\nu}dx^\mu dx^\nu$. Answer:

$$dS^2 = e^{2\xi\tilde{a}/c^2}\left(-c^2 d\lambda^2 + d\xi^2\right)$$

Problem 85: Using the equations of motion (163), determine the alternative Rindler metric again by applying the tensor transformation law. Answer:

$$g_{\mu\nu} = \frac{\partial\tilde{x}^\rho}{\partial x^\mu}\frac{\partial\tilde{x}^\lambda}{\partial x^\nu}\eta_{\rho\lambda}$$

$$g_{\lambda\lambda} = -e^{2\xi\tilde{a}/c^2} \qquad\qquad g_{\lambda\xi} = g_{\xi\lambda} = 0 \qquad\qquad g_{\xi\xi} = e^{2\xi\tilde{a}/c^2}$$

Problem 86: The Rindler *time translation vector* is defined as $V = \partial/\partial\lambda$, and has components

$$V^\lambda = 1 \qquad\qquad V^\xi = 0\ .$$

Express the components of V in Cartesian coordinates. Answer:

$$V^t = x\frac{\tilde{a}}{c} \qquad\qquad V^x = t\frac{\tilde{a}}{c}$$

8.5 Rindler Diagram

Recall the equations of motion (156), (157) and separately consider lines of constant τ and ξ. For lines of constant τ, the equations of motion combine to give

$$x + \frac{c^2}{\tilde{a}} = \coth\left(\tilde{a}\tau/c\right)\cdot ct\ ,$$

where the coth() function is a dimensionless number depending on τ. Clearly these are straight lines in spacetime. Meanwhile, lines of constant ξ imply the relation

$$-\frac{(ct)^2}{\left(c^2/\tilde{a}+\xi\right)^2} + \frac{\left(x + c^2/\tilde{a}\right)^2}{\left(c^2/\tilde{a}+\xi\right)^2} = 1\ ,$$

which is the equation of a hyperbola in spacetime. Plotting families of these lines onto one spacetime diagram, we get a picture called a *Rindler diagram*.

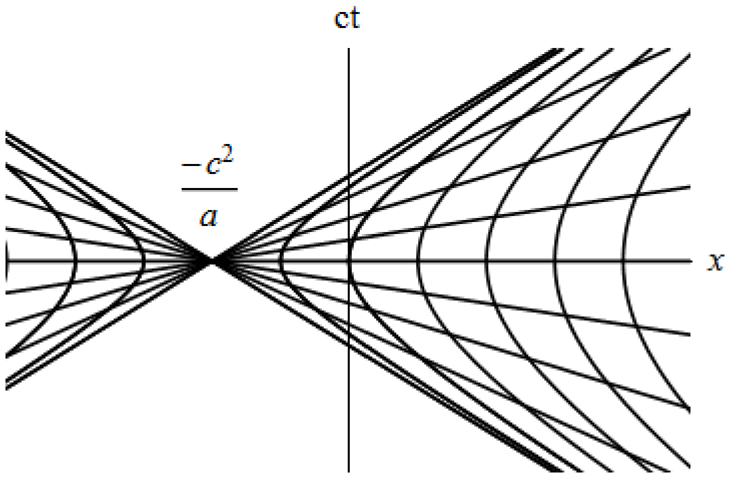

Figure 2: Rindler diagram with curves of constant τ and coonstant ξ.

The Rindler diagram is a visual reminder of the 'weirdness' of Rindler space. Objects of constant ξ clearly don't trace out straight lines, hence Rindler space is considered *curved*. The time τ is represented by straight lines, which are initially horizontal and asymptotically evolve toward lines of slope 45° as τ increases.

The lines of constant τ intersect at $x = -c^2/\tilde{a}$. The equations of motion are non-analytic across this point, implying that only half of Rindler space is physically available (in a rocket, for example). We by convention choose to throw away the space where $x < -c^2/\tilde{a}$, which has curious implications. Chiefly, we wee that Rindler space covers only a quater of Minkowski space. Physically, this means that information originating in the 'forbidden zone' (white space or any points left of $-c^2/\tilde{a}$) can *never* penetrate or influence Rindler space. (Certain signals will never ever reach the rocket.)

9 Geometry in Curved Space

The ordinary rules of geometry are heavily used in physics due to their robustness: it's difficult to conceive of a right triangle that does not obey the pythagorean theorem, or to draw a circle whose diameter is not proportional to the radius. Such things are impossible in ordinary Euclidean or 'flat' space.

Consider a bug that is confined to live on the surface of a large ball, and the bug's mission is to find out whether or not he lives on a curved space. Only movements *on* the surface are allowed, so he's not free to rise up and look for a curving horizon. A bug that has studied Euclid knows that any triangle drawn on a plane should have precisely 180 degrees as the sum of its interior angles. Trying this, the bug on a beachball will find that the sum is actually *greater* than 180 degrees. Right away he concludes that the surface is not planar. (Had be been confined to a horse saddle instead of a ball, the sum would be less than 180.)

Therefore, let's loosely define *curved space* as any space where the ordinary rules of geometry don't work. The pressing question is the, what *are* the rules of geometry for curved sapce?

9.1 Basis Vectors, Metric, Tangent Vectors

A space that is completely covered in a differentiable and locally flat coordinate system is called a *manifold*. For general relativity (our case), a manifold must also contain a time dimension.

Let the position vector on a manifold be denoted $\vec{r}$, which is a function of parameters q^1, q^2, q^3, ..., q^N, with N being the number of dimensions on the manifold. The units of any q^j are not limited to length. By the chain rule, the differential line element on a manifold is

$$d\vec{S} = \frac{\partial \vec{r}}{\partial q^1} dq^1 + \frac{\partial \vec{r}}{\partial q^2} dq^2 + \cdots \ . \tag{164}$$

The terms $\partial \vec{r}/\partial q^j$ are interpreted as basis vectors on the manifold. You'll notice these are variable with $\vec{r}$ and are not normalized. Denoting each basis vector $\vec{a}_{(j)}$, we write

$$d\vec{S} = \sum \vec{a}_{(\mu)} dq^\mu \ . \tag{165}$$

The space(time) interval dS^2 is equal to the square of the differential line element, coming out to

$$dS^2 = \sum\sum \left(\vec{a}_{(\mu)} \cdot \vec{a}_{(\nu)}\right) dq^\mu dq^\nu \ .$$

The metric of a curved coordinate system shall be denoted $g_{\mu\nu}$, which is the generalization of the flat metric $\eta_{\mu\nu}$. In our new construction, the metric is defined as

$$g_{\mu\nu} = \vec{a}_{(\mu)} \cdot \vec{a}_{(\nu)} \ , \tag{166}$$

giving

$$dS^2 = g_{\mu\nu} dq^\mu dq^\nu \ , \tag{167}$$

where the Einstein summation convention has been used. Components of $g_{\mu\nu}$ obey the orthogonality relation

$$g^{\mu\alpha} g_{\alpha\nu} = \vec{a}^{(\mu)} \cdot \vec{a}_{(\nu)} = \delta^\mu_\nu \ . \tag{168}$$

Any tangent vector $\vec{V}$ that lives on a manifold must admit an expansion in terms of basis vectors, as in

$$\vec{V} = V^\mu \vec{a}_{(\mu)} \ . \tag{169}$$

Note (again) that the basis vectors are not normalized. That is, you are probably more familiar with a unit basis, as in

$$\vec{V} = v^\mu \hat{e}_{(\mu)} \ .$$

Don't let the difference confuse you!

The definition of the four-velocity needs some alteration for curved space. Recall first that the spacetime interval dS^2 and the proper time interval relate by $dS^2 = -c^2 d\tau^2$. Mentally dividing equation (167) through by $d\tau^2$, we see that the velocity four-vector should be redefined as

$$U^\mu = \frac{\partial q^\mu}{\partial \tau} \tag{170}$$

in order to obey the normalization condition $U^\mu U_\mu = -c^2$.

9.2 Common Coordinate Systems

Now build a catalog of common metrics $g_{\mu\nu}$. In the results that follow, the time component g^{00} is implied if inot explicitly mentioned.

Cartesian: 2D

The trivial representation of a planar space is $\vec{r} = x\hat{x} + y\hat{y}$, implying $d\vec{r} = dx\hat{x} + dy\hat{y}$. The square of a differential line element is equal to

$$dS^2 = \vec{dr} \cdot \vec{dr} = dx^2 + dy^2 \, ,$$

which shouldn't be startling (it's just the pythagorean theorem). For now on the phrase 'square of the differntial line element' shall be synonymous with 'metric'.

Polar: 2D

Suppose we choose a different parameterization to cover the xy-plane, namely

$$x = r\cos\theta \qquad\qquad y = r\sin\theta \, , \tag{171}$$

with the inverted version

$$r = \sqrt{x^2 + y^2} \qquad\qquad \theta = \tan^{-1}\left(\frac{y}{x}\right) \, .$$

The metric is obviously

$$dS^2 = dr^2 + r^2 d\theta^2 \, , \tag{172}$$

where comparison to (167) tells us $dq^r = dr$, $dq^\theta = d\theta$, and meanwhile $dq_r = dr$, $dq_\theta = r^2 d\theta$.

Problem 87: Express the $2D$ plane polar metric (172) in tensor component form. That is, find g_{rr}, $g_{\theta\theta}$, g^{rr}, and $g^{\theta\theta}$. Answer:

$$g_{rr} = 1 \qquad\qquad g_{\theta\theta} = r^2$$

$$g^{rr} = 1 \qquad\qquad g^{\theta\theta} = \frac{1}{r^2}$$

Problem 88: Suppose a vector $\vec{V}$ is based at the point $x = 1/\sqrt{2}$, $y = 1/\sqrt{2}$, and has components $V^x = 3$, $V^y = 1$. Compute the components of $\vec{V}$ in polar coordinates. Answer:

$$\tilde{V}^\mu = \frac{\partial \tilde{q}^\mu}{\partial q^\nu} V^\nu \qquad\qquad V^r = 2\sqrt{2} \qquad\qquad V^\theta = \sqrt{2}$$

(What you don't need to know is $|V| = \sqrt{10}$, or that the angle between the vector $\vec{V}$ and the vector $\vec{r}_0$ locating its base is $\approx -27°$.)

Parabolic: 2D

Consider the coordinate transformation from $2D$ Cartesian coordinates (x, y) to the new set of coordinates (u, v),

$$x = uv \qquad\qquad y = \frac{1}{2}\left(u^2 - v^2\right) \, , \tag{173}$$

known as *parabolic coordinates.*

Problem 89: Show that the metric for parabolic coordinates is

$$dS^2 = \left(v^2 + v^2\right)\left(du^2 + dv^2\right) \, .$$

Problem 90: Find the equation of a circle of radius R centered at the origin in terms of coordinates (u, v). Answer:

$$u^2 + v^2 = 2R$$

Cartesian: 3D

Any point in three-dimensional space is located at Cartesian coordinates $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$. The square of the differential line element $d\vec{r} \cdot d\vec{r}$ is nothing more than the spatial part of the Minkowski metric:

$$dS^2 = dx^2 + dy^2 + dz^2$$

Cylindrical: 3D

Problem 91: Write down the metric for flat space in cylindrical coordinates. Answer:

$$dS^2 = dr^2 + r^2 d\phi^2 + dz^2 \tag{174}$$

Spherical: 3D

Now let us locate points in three-dimensional space by spherical coordinates, namely

$$x = r \sin\theta \cos\phi \qquad y = r \sin\theta \sin\phi \qquad z = r\cos\theta \,. \tag{175}$$

Of course we could compute the metric dS^2 by brute-force substitution of (175) into $\sum dx_i^2$ (which you are encouraged to do once), but please get accustomed to using the transformation law (61), which applies to our problem as follows:

$$g_{\alpha\beta} = \frac{\partial q^\sigma}{\partial \tilde{q}^\alpha} \frac{\partial q^\rho}{\partial \tilde{q}^\beta} \eta_{\sigma\rho}$$

The components of $g_{\alpha\beta}$ turn out to be

$$g_{rr} = 1 \qquad g_{\theta\theta} = r^2 \qquad g_{\phi\phi} = r^2 \sin^2\theta \,,$$

so the metric in spherical coordinates is

$$dS^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \,. \tag{176}$$

Note the above doesn't look like the Cartesian flat metric tensor, but indeed describes flat space.

A third way to derive equation (176) is to write down the line element in spherical coordinates, namely

$$d\vec{S} = dr\hat{r} + rd\theta\hat{\theta} + r\sin\theta d\phi\hat{\phi} \,. \tag{177}$$

To determine $g_{\mu\nu}$, take $d\vec{S} \cdot d\vec{S}$ and read off the coefficients on dr^2, $d\theta^2$, and $d\phi^2$.

Problem 92: For $3D$ spherical coordinates, write down the basis vectors $\vec{a}_{(r)}$, $\vec{a}_{(\theta)}$, $\vec{a}_{(\phi)}$ and their inverses $\vec{a}^{(r)}$, $\vec{a}^{(\theta)}$, $\vec{a}^{(\phi)}$. Answer:

$$\vec{a}_{(r)} = \hat{r} \qquad \vec{a}_{(\theta)} = r\hat{\theta} \qquad \vec{a}_{(\phi)} = r\sin\theta\hat{\phi}$$

$$\vec{a}^{(r)} = \hat{r} \qquad \vec{a}^{(\theta)} = \frac{\hat{\theta}}{r} \qquad \vec{a}^{(\phi)} = \frac{\hat{\phi}}{r\sin\theta}$$

Spherical: 2D

A special case of spherical coordinates occurs when the radius is fixed, in which case the available space has only two dimensions. The metric on a spherical surface is

$$dS^2 = R^2 \left(d\theta^2 + \sin^2\theta d\phi^2 \right) \,. \tag{178}$$

Nonround sphere: 2D

An axisymmetric, but not necessarily round, sphere has a line element of the form

$$dS^2 = R^2 \left(d\theta^2 + f^2(\theta)\, d\phi^2\right) \,, \tag{179}$$

where $f(\theta)$ is equal to $\sin\theta$ for a round sphere.

Problem 93: The Earth is not a perfectly round sphere. Instead, the polar radius, measured to be $6357\,km$, is slightly less than the equatorial radius, measured as $6378\,km$. Supposing the surface of the earth is modeled by

$$f(\theta) = \sin\theta \left(1 + \epsilon \sin^2\theta\right) \,,$$

what values of R and ϵ would best represent the data? Answer:

$$R = 6557\,km \qquad\qquad \epsilon = 0.0033$$

Spherical: 4D

The three-sphere is a higher-dimensional analog of the ordinary sphere. In the sense that an ordinary sphere is defined as the $2D$ set of points that lie a fixed distance from a common center, a three-sphere is a $3D$ surface embedded in $4D$ space. A three-sphere can be expressed in *hyperspherical coordinates* as:

$$x = r\cos\psi \qquad\qquad y = r\sin\psi\cos\theta$$

$$z = r\sin\psi\sin\theta\cos\phi \qquad\qquad w = r\sin\psi\sin\theta\sin\phi$$

Problem 94: Define a four-dimensional flat space metric $\eta_{\mu\nu}$ as diag$\{1,1,1,1\}$. Using the relation

$$g_{\alpha\beta} = \frac{\partial q^\sigma}{\partial \tilde{q}^\alpha}\frac{\partial q^\rho}{\partial \tilde{q}^\beta}\eta_{\sigma\rho} \,,$$

show that the metric on the three-sphere of unit radius is

$$dS^2 = d\psi^2 + \sin^2\psi d\theta^2 + \sin^2\psi\sin^2\theta d\phi^2 \,. \tag{180}$$

Something not obvious is whether a given metric represents flat space or curved space. While we don't yet have the tools to decide this for equation (180), you can safely assume that the three-sphere is a curved space.

Rindler space: 1+3D

We wrote down the metric for $1{+}1D$ Rinder space as equation (158). The spatial dimension ξ is along the direction of uniform acceleration. Generalizing to three dimension, one might write

$$dS^2 = -c^2 d\tau^2 \left(1 + \frac{\xi_x \tilde{a}}{c^2}\right)^2 + d\xi_x^2 + d\xi_y^2 + d\xi_z^2 \,.$$

Milne space: 1+1D

The Milne spacetime is a toy model of an expanding universe. Starting with the $1+1D$ Minkowski metric $dS^2 = -c^2dt^2 + dx^2$ and transforming to the coordinates (T, z), we write:

$$t = T\cosh kz \qquad\qquad x = T\sinh kz$$

Problem 95: Show that the metric in Milne space time is given by

$$dS^2 = -c^2dT^2 + c^2T^2k^2dz^2\ .$$

Friedman-Robertson-Walker space: 1+3D

Friedman-Robertson-Walker space is homogeneously expanding (or contracting) Minkowski space. The metric in FRW space is given by

$$dS^2 = -c^2dt^2 + a^2\left(t\right)\delta_{kl}dx^kdx^l\ , \tag{181}$$

where $a(t)$ is a general function of time.

9.3 Gradient in Curved Space

The dx Notation

The four-gradient was previously introduced in equation (63), stating that for a scalar field $f(\vec{r})$, the components are

$$\partial_\mu f = \left(\frac{\partial f}{\partial x^0}, \frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^N}\right)\ .$$

Note that $\partial_\mu \phi$ is only written in component form. Converting to full-blown vector notation, we have

$$\vec{\nabla} f = \hat{a}^{(\mu)}\frac{\partial f}{\partial x^\mu} \tag{182}$$

Testing this construction on polar coordinates as an example, identify

$$dx^r = dr \qquad\qquad dx^\theta = rd\theta$$

for differentials,

$$\vec{a}_{(r)} = \hat{r} \qquad\qquad \vec{a}_{(\theta)} = r\hat{\theta}$$

$$\vec{a}^{(r)} = \hat{r} \qquad\qquad \vec{a}^{(\theta)} = r^{-1}\hat{\theta}$$

for basis vectors, and for the gradient,

$$\vec{\nabla} f = \frac{\partial f}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\hat{\theta}\ .$$

The dq Notation

The gradient operator (182) can be rewritten as

$$\vec{\nabla} f = \vec{a}^{(\mu)} \frac{\partial f}{\partial q^\mu} \,, \tag{183}$$

because

$$\vec{\nabla} f = \hat{a}^{(\mu)} \frac{\partial q^\alpha}{\partial x^\mu} \frac{\partial f}{\partial q^\alpha} = \frac{\hat{a}^{(\alpha)}}{\left|\vec{a}_{(\mu)}\right|} \frac{\partial f}{\partial q^\alpha} = \frac{\vec{a}^{(\alpha)}}{\left|\vec{a}_{(\mu)}\right| \left|\vec{a}^{(\alpha)}\right|} \frac{\partial f}{\partial q^\alpha} = \vec{a}^{(\mu)} \frac{\partial f}{\partial q^\mu} \,.$$

Repeating the previous exercise in polar coordinates, we now have

$$dq^r = dr \qquad\qquad dq^\theta = d\theta \,,$$

where the basis vectors $\vec{a}_{(\mu)}$ are unchanged. The gradient now trivially falls out of equation (183).

Remark:

The dx notation is old news, and is only finitely useful. To proceed in general, favor the dq notation.

Problem 96: Take the function $f(x, y) = \ln \sqrt{x^2 + y^2}$ and calculate components of the gradient $\vec{\nabla} f$ in Cartesian coordinates, and write down the components V^x, V^y. Next repeat the exercise in polar coordinates by two different methods: (i) substitute x and y into f to make $f(r)$; (ii) use the vector transformation law on the result attained for Cartesian coordinates. Answer:

$$V^x = \frac{x}{r^2} \qquad\qquad V^y = \frac{y}{r^2}$$

$$\tilde{V}^\mu = \frac{\partial \tilde{q}^\mu}{\partial q^\nu} V^\nu \qquad\qquad V^r = \frac{1}{r} \qquad\qquad V^\theta = 0$$

Problem 97: Use equation (183) to derive the formula for the gradient of a scalar field $f(r, \theta, \phi)$ in $3D$ spherical coordinates. Answer:

$$\vec{\nabla} f = \frac{\partial f}{\partial r} \hat{r} + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\theta} + \frac{1}{r \sin\theta} \frac{\partial f}{\partial \phi} \hat{\phi} \tag{184}$$

9.4 Covariant Derivative and the Connection

We've seen that the four-gradinet converts a scalar field $\phi(x^\mu)$ into a covariant vector field, for instance $\partial_\mu \phi = V_\mu$. We now may wonder, is the four-divergence $\partial_\mu \partial_\nu \phi$ actually a tensor? We find out by looking at coordinate transformations, which we *hope* look like

$$\partial_{\mu'} \partial_{\nu'} \phi \stackrel{?}{=} \frac{\partial q^\alpha}{\partial q^{\mu'}} \frac{\partial q^\beta}{\partial q^{\nu'}} \partial_\alpha \partial_\beta \phi \,.$$

Checking for this, we instead find

$$\partial_{\mu'} \partial_{\nu'} \phi = \frac{\partial q^\alpha}{\partial q^{\mu'}} \frac{\partial q^\beta}{\partial q^{\nu'}} \partial_\alpha \partial_\beta \phi + \frac{\partial q^\alpha}{\partial q^{\mu'}} \frac{\partial^2 q^\beta}{\partial q^\alpha \partial q^{\nu'}} \frac{\partial \phi}{\partial q^\beta} \,, \tag{185}$$

and there is an ugly extra term on the right side. Evidently $\partial_\mu \partial_\nu \phi$ does not transform as a tensor. (It looks like a duck but does not quack.)

Equivalently, the derivative (with respect to q^μ, not x^μ) of a contravariant tangent vector field V^μ suffers the same problem because the basis vectors $\vec{a}_{(\mu)}$ have their own derivatives:

$$\partial_\nu \left(V^\mu \vec{a}_{(\mu)} \right) = \vec{a}_{(\mu)} \left(\partial_\nu V^\mu \right) + V^\mu \partial_\nu \vec{a}_{(\mu)}$$

The latter term in the above involves components that stick off of the manifold, meaning

$$V^\mu \partial_\mu \vec{a}_{(\mu)} = \vec{T} + \vec{N} \; ,$$

where T and N stand for 'tangential' and 'normal' components, respectively. We want vectors, tensors, and their derivatives to always live on the manifold, so the idea is to subtract away the unwanted normal component from the 'definition' of the derivative. In doing so we introduce the *covariant derivative operator*, denoted by uppercase D, as

$$D_\nu \left(V^\mu \vec{a}_{(\mu)} \right) = \partial_\nu \left(V^\mu \vec{a}_{(\mu)} \right) - \vec{N} \; ,$$

meaning

$$D_\nu \left(V^\mu \vec{a}_{(\mu)} \right) = \vec{a}_{(\mu)} \left(\partial_\nu V^\mu \right) + \Gamma^\mu_{\nu\beta} V^\beta \vec{a}_{(\mu)} \; . \tag{186}$$

The tangential component $\vec{T}$ has been exressed in terms of basis vectors $\vec{a}_{(\mu)}$ with unknown coeffcients contained in the symbols $\Gamma^\mu_{\nu\beta}$, known as *Christoffel symbols* or *connection coefficients*. It's important to note now that $\Gamma^\mu_{\nu\beta}$ is *not* a tensor, as we'll prove later. In component form, the covariant derivate of V^μ reads

$$D_\nu V^\mu = \partial_\nu V^\mu + \Gamma^\mu_{\nu\beta} V^\beta \; , \tag{187}$$

one of the most frequently employed equations in this entire subject.

Properties of the Covariant Derivative

Like the ordinary derivative, the covariant derivative satisfies

$$D_\alpha \left(U^\mu + V^\mu \right) = D_\alpha U^\mu + D_\alpha V^\mu$$

and

$$D_\alpha \left(U^\mu V^\nu \right) = \left(D_\alpha U^\mu \right) V^\nu + U^\mu \left(D_\alpha V^\nu \right) \; .$$

Another important feature of D_μ is the commutation with contractions:

$$D_\mu \left(A^\alpha_{\alpha\gamma} \right) = \left(DA \right)^\alpha_{\mu\alpha\gamma}$$

Without worrying (yet) about the components of Γ, equation (186) is readily generalized to calculate derivatives of any object. For scalar fields, the covariant derivative is simply partial differentiation:

$$D_\alpha \phi = \partial_\alpha \phi \tag{188}$$

For covariant tangent vector fields V_μ, it turns out that

$$D_\nu V_\mu = \partial_\nu V_\mu - \Gamma^\rho_{\nu\mu} V_\rho \; , \tag{189}$$

where there now is a minus sign ahead of the symbol Γ. Equations (186) and (189) tell us that that D_α acting on V with an 'up' index involves $+\Gamma$, and for a 'down' index we need $-\Gamma$. Covariant derivatives of higher-order objects introduce one Γ term for each index. To illustrate, one might encounter any of

$$D_\alpha T^{\mu\nu} = \partial_\alpha T^{\mu\nu} + \Gamma^\mu_{\alpha\beta} T^{\beta\nu} + \Gamma^\nu_{\alpha\rho} T^{\mu\rho}$$

$$D_\alpha U_{\mu\nu} = \partial_\alpha U_{\mu\nu} - \Gamma^\rho_{\alpha\mu} U_{\rho\nu} - \Gamma^\gamma_{\alpha\nu} U_{\mu\gamma}$$

$$D_\alpha V^\mu_\nu = \partial_\alpha V^\mu_\nu + \Gamma^\mu_{\alpha\gamma} V^\gamma_\nu - \Gamma^\beta_{\alpha\nu} V^\mu_\beta \ ,$$

where the V^μ_ν equation may be used to verify (189) by setting $\mu = \nu$.

Now we prove why the minus signs belong in front of certain Γ. Supposing we didn't know equation (189) yet, one might guess that

$$D_\nu V_\mu = \partial_\nu V_\mu + \bar{\Gamma}^\rho_{\nu\mu} V_\rho \ ,$$

where $\bar{\Gamma}$ has something to do with the original Γ. The relationship is found by writing out the covariant derivative D_μ of a scalar field $\phi = V^\alpha U_\alpha$ in two ways:

$$D_\mu \phi = \partial_\mu \phi = \partial_\mu \left(V^\alpha U_\alpha \right)$$

$$D_\mu \phi = U_\alpha \left(\partial_\mu V^\alpha + \Gamma^\alpha_{\mu\beta} V^\beta \right) + V^\alpha \left(\partial_\mu U_\alpha + \bar{\Gamma}^\rho_{\mu\alpha} U_\rho \right)$$

Comparing the above, we notice $0 = U_\gamma V^\alpha \left(\Gamma^\gamma_{\mu\alpha} + \bar{\Gamma}^\gamma_{\mu\alpha} \right)$, which can only be true if

$$\Gamma^\gamma_{\mu\alpha} = -\bar{\Gamma}^\gamma_{\mu\alpha} \ ,$$

verifying the ansatz (189).

Coordinate Transformations

The object $\Gamma^\mu_{\nu\beta}$ is not itself a tensor, which is verified by examining Γ under coordinate transformations. Fortunately, covariant derivatives such as $D_\nu V^\mu$ are indeed tensors, which under coordinate transformations appear as

$$D_{\nu'} V^{\mu'} = \frac{\partial q^\alpha}{\partial q^{\nu'}} \frac{\partial q^{\mu'}}{\partial q^\beta} D_\alpha V^\beta \ ,$$

which could have instead been written

$$D_{\nu'} V^{\mu'} = \frac{\partial q^\alpha}{\partial q^{\nu'}} \frac{\partial}{\partial q^\alpha} \left(\frac{\partial q^{\mu'}}{\partial q^\beta} V^\beta \right) + \tilde{\Gamma}^{\mu'}_{\nu'\rho'} \left(\frac{\partial q^{\rho'}}{\partial q^\beta} V^\beta \right) \ .$$

Eliminating $D_{\nu'} V^{\mu'}$, arrive at (after a cancellation and shuffling a few indices)

$$\frac{\partial q^\alpha}{\partial q^{\nu'}} \frac{\partial q^{\mu'}}{\partial q^\beta} \Gamma^\beta_{\alpha\gamma} V^\gamma = \frac{\partial q^\alpha}{\partial \nu'} \frac{\partial^2 q^{\mu'}}{\partial q^\alpha \partial q^\gamma} V^\gamma + \tilde{\Gamma}^{\mu'}_{\nu'\rho'} \frac{\partial q^{\rho'}}{\partial q^\gamma} V^\gamma \ ,$$

which must be true for all V^γ. Conclude

$$\tilde{\Gamma}^{\mu'}_{\nu'\rho'} = \frac{\partial q^\alpha}{\partial q^{\nu'}} \frac{\partial q^{\mu'}}{\partial q^\beta} \frac{\partial q^\gamma}{\partial q^{\rho'}} \Gamma^\beta_{\alpha\gamma} - \frac{\partial q^\alpha}{\partial \nu'} \frac{\partial^2 q^{\mu'}}{\partial q^\alpha \partial q^\gamma} \frac{\partial q^\gamma}{\partial q^{\rho'}} \ , \tag{190}$$

sealing the deal that Γ is not a tensor (the second term wouldn't be there otherwise).

9.5 Torsion Tensor

A combination of Christoffel symbols actually can constitute a tensor. The quantity

$$T^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{\lambda\nu} - \Gamma^{\mu}_{\nu\lambda} \tag{191}$$

is called the *torsion*. Of course, we don't know for sure whether $T^{\mu}_{\nu\lambda}$ qualifies as a tensor until we try to calculate it in another coordinate system, which you are encouraged to check. Given equation (190), you can see $\tilde{T}^{\mu'}_{\nu'\lambda'}$ subtracts all non-tensorlike terms away.

In the theory of general relativity, the torsion is assumed to be zero. By construction, if a tensor is zero in some reference frame, then it's zero in all reference frames. Thus the torsion is zero in all contexts, and we adopt an important identity:

$$\Gamma^{\mu}_{\lambda\nu} = \Gamma^{\mu}_{\nu\lambda} \tag{192}$$

You may wonder how torsion could be anything other than zero. The phenomenon of quantum spin has caused physicists to think twice about the universe having zero torsion, but for all purposes we shall take $T^{\mu}_{\nu\lambda} = 0$. Equation (192) tells us the Christoffel symbols are symmetric in the two lower indices. Still though, there are 40 nontrivial components in $\Gamma^{\mu}_{\lambda\nu}$.

Problem 98: Show that if $A_{\mu\nu} = -A_{\nu\mu}$ is an antisymmetric type $(0,2)$ tensor, all Christoffel symbols cancel out of

$$D\left[_{\mu}A_{\nu\rho}\right] = \partial\left[_{\mu}A_{\nu\rho}\right] \; .$$

9.6 Metric Compatibility

So far we've written few restrictions on the Christoffel symbols $\Gamma^{\mu}_{\nu\rho}$, particularly in their relation to the metric $g_{\alpha\beta}$. With the leftover freedom, let's require that the condtion

$$D_{\alpha} g_{\mu\nu} = 0 \tag{193}$$

always holds true. This is known as the *metric compatibility condition*. This is a good move because one is allowed to raise and lower indices through the directional derivative. Given (193), the inverse metric also obeys the compatibility condition, namely

$$D_{\alpha} g^{\mu\nu} = 0 \; .$$

9.7 Christoffel Symbols

The metric compatibility condtion shall be used to work out an extremely useful formula: a way to determine the connection coefficients $\Gamma^{\mu}_{\nu\rho}$ from the metric $g_{\alpha\beta}$. Writing out the covariant derivative in equation (193), we have

$$D_{\alpha} g_{\mu\nu} = \partial_{\alpha} g_{\mu\nu} - \Gamma^{\gamma}_{\alpha\mu} g_{\gamma\nu} - \Gamma^{\gamma}_{\alpha\nu} g_{\mu\gamma} \; .$$

The indices α, μ, ν can be permutated twice to yield two more equivalent statements

$$D_{\nu} g_{\alpha\mu} = \partial_{\nu} g_{\alpha\mu} - \Gamma^{\gamma}_{\nu\alpha} g_{\gamma\mu} - \Gamma^{\gamma}_{\nu\mu} g_{\alpha\gamma} \; ,$$

$$D_{\mu} g_{\nu\alpha} = \partial_{\mu} g_{\nu\alpha} - \Gamma^{\gamma}_{\mu\nu} g_{\gamma\alpha} - \Gamma^{\gamma}_{\mu\alpha} g_{\nu\gamma} \; .$$

Subtracting the second two equations from the first, and also taking advantage of the symmetry $\Gamma^{\alpha}_{\beta\gamma} = \Gamma^{\alpha}_{\gamma\beta}$, arrive at

$$0 = \partial_\alpha g_{\mu\nu} - \partial_\nu g_{\alpha\mu} - \partial_\mu g_{\mu\alpha} + 2\Gamma^{\gamma}_{\mu\nu} g_{\gamma\alpha} \,,$$

which amazingly has just one instance of the Christoffel symbol Γ. Isolating this, we finally discover

$$\Gamma^{\beta}_{\mu\nu} = \frac{1}{2} g^{\beta\alpha} \left(\partial_\mu g_{\nu\alpha} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu} \right) \,, \tag{194}$$

something worth memorizing.

In Minkowski space, and also in $2D$ and $3D$ Cartesian coordinate systems, all connection coefficients Γ are zero for obvious reasons. This does *not* to say that all $\Gamma^{\beta}_{\mu\nu} = 0$ for flat space - the choice of coordinates determine Γ, which alone doesn't indicate whether a particular space is curved.

Problem 99: Apply equation (194) to the metric (172) to determine all nonzero Christoffel symbols for $2D$ plane polar coordinates. Answer:

$$\Gamma^{\theta}_{r\theta} = \frac{1}{r} \qquad \Gamma^{\theta}_{\theta r} = \frac{1}{r} \qquad \Gamma^{r}_{\theta\theta} = -r$$

Problem 100: Apply equation (194) to the metric (174) to determine all nonzero Christoffel symbols for $3D$ cylindrical coordinates. Answer:

$$\Gamma^{\phi}_{r\phi} = \frac{1}{r} \qquad \Gamma^{\phi}_{\phi r} = \frac{1}{r} \qquad \Gamma^{r}_{\phi\phi} = -r$$

Problem 101: Apply equation (194) to the metric (176) to determine all nonzero Christoffel symbols for $3D$ spherical coordinates. Answer:

$$\Gamma^{\theta}_{r\theta} = \Gamma^{\phi}_{r\phi} = \frac{1}{r} \qquad \Gamma^{r}_{\theta\theta} = -r \qquad \Gamma^{r}_{\phi\phi} = -r\sin^2\phi$$

$$\Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta \qquad \Gamma^{\phi}_{\phi\theta} = \cot\theta$$

Problem 102: Apply equation (194) to the metric (178) to determine all nonzero Christoffel symbols for a two-sphere (one which has fixed radius embedded in three-dimensional space). Answer:

$$\Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta \qquad \Gamma^{\phi}_{\phi\theta} = \cot\theta$$

Problem 103: We've seen that the metric for the three-sphere of fixed radius in $4D$ is contained in equation (180). Choosing a sligtly different parameterization, it's possible to derive the metric

$$dS^2 = d\theta^2 + \sin^2\theta d\psi^2 + \cos^2\theta d\phi^2$$

for the same space. Apply equation (194) to to determine all nonzero Christoffel symbols for the metric given. Answer:

$$\Gamma^{\theta}_{\psi\psi} = -\sin\theta\cos\theta \qquad \Gamma^{\theta}_{\phi\phi} = \sin\theta\cos\theta$$

$$\Gamma^{\psi}_{\theta\psi} = \cos\theta / \sin\theta \qquad \Gamma^{\phi}_{\theta\phi} = -\sin\theta / \cos\theta$$

Problem 104: Use the Rindler metric (160) to determine all nonzero Christoffel symbols for Rindler space. Answer:

$$\Gamma^{\xi}_{\tau\tau} = \frac{\tilde{a}}{c^2}\left(1 + \frac{\xi\tilde{a}}{c^2}\right) \qquad \Gamma^{\tau}_{\xi\tau} = \frac{\tilde{a}/c^2}{1 + \xi\tilde{a}/c^2}$$

Problem 105: Apply equation (194) to the Friedman-Robertson-Walker metric (181) to determine all nonzero Christoffel symbols for FRW space. Answer:

$$\Gamma^t_{kk} = a\dot{a} \qquad \Gamma^k_{tk} = \frac{\dot{a}}{a} \tag{195}$$

9.8 Divergence, Curl, Laplacian Operators

The covariant derivative equation (186) can be used to deliver a flurry of useful identities. Multiplying through by the basis vector $\vec{a}^{(\nu)}$ recovers the formula for the divergence of a vector:

$$\vec{\nabla} \cdot \vec{V} = \left(\vec{a}^{(\nu)} \cdot \vec{a}_{(\mu)}\right) (D_\nu V^\mu) \tag{196}$$

Similarly, the curl of a vector $\vec{V}$ in curved space is given by

$$\vec{\nabla} \times \vec{V} = \left(\vec{a}^{(\nu)} \times \vec{a}_{(\mu)}\right) (D_\nu V^\mu) \ . \tag{197}$$

The Laplacian operator $\nabla^2 = \Delta$ is naturally expressed in terms of covariant derivatives:

$$\nabla^2 \phi = g^{\mu\nu} D_\mu D_\nu \phi \tag{198}$$

Problem 106: Use equation (196) to calculate the divergence of a vector $\vec{V} = V^\mu \vec{a}_{(\mu)} = v_r \hat{r} + v_\theta \hat{\theta} + v_\phi \hat{\phi}$ in $3D$ spherical coordinates. Answer:

$$\begin{aligned}
\vec{\nabla} \cdot \vec{V} &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 V^r\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta V^\theta\right) + \partial_\phi V^\phi \\
&= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 v_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta v_\theta\right) + \frac{1}{r\sin\theta}\partial_\phi v_\phi
\end{aligned}$$

Problem 107: Use equation (197) to calculate the curl of a vector $\vec{V} = V^\mu \vec{a}_{(\mu)} = v_r \hat{r} + v_\theta \hat{\theta} + v_\phi \hat{\phi}$ in $3D$ spherical coordinates. Answer:

$$\begin{aligned}
\left(\vec{\nabla} \times \vec{V}\right)_r &= \sin\theta\partial_\theta V^\phi + 2\cos\theta V^\phi - \frac{\partial_\phi V^\theta}{\sin\theta} \\
&= \frac{1}{r\sin\theta}\left(\partial_\theta\left(\sin\theta v_\phi\right) - \partial_\phi v_\theta\right) \\
\left(\vec{\nabla} \times \vec{V}\right)_\theta &= -r\sin\theta\partial_r V^\phi - 2\sin\theta V^\phi + \frac{\partial_\phi V^r}{r\sin\theta} \\
&= \frac{1}{r}\left(\frac{\partial_\phi v_r}{\sin\theta} - \partial_r\left(r v_r\right)\right) \\
\left(\vec{\nabla} \times \vec{V}\right)_\phi &= r\partial_r V^\theta + 2V^\theta - \frac{\partial_\theta V^r}{r} \\
&= \frac{1}{r}\left(\partial_r\left(r v_\theta\right) - \partial_\theta v_r\right)
\end{aligned}$$

Problem 108: Use the Laplacian operator (198) to calculate the $\Delta\phi(\vec{r})$ on a sphere of fixed radius embedded in three-dimensional space. Answer:

$$\nabla^2\phi = g^{\mu\nu} D_\mu D_\nu \phi = \frac{1}{R^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{R^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}$$

9.9 Parallel Transport

The Christoffel symbols, a.k.a. connection coefficients $\Gamma^{\beta}_{\mu\nu}$ were summoned into being by requiring that covariant derivatives D_α of a tangent vector field $V^\mu(x^\mu)$ stay on the manifold in which V^μ is living. Here we develop a more 'physical' interpretation of the connection coefficients by introducing the notion of *parallel transport* of a vector. As the name suggests, parallel transport is the translation (without rotations) of a vector on a manifold.

Consider a $2D$ example in plane polar coordinates: suppose we have a vector V^μ with components V^r, V^θ that is based at the point $\vec{X} = (r\cos\theta, r\sin\theta)$. The vector V makes an angle ϕ with respect to the position vector X as shown in figure Figure 3.

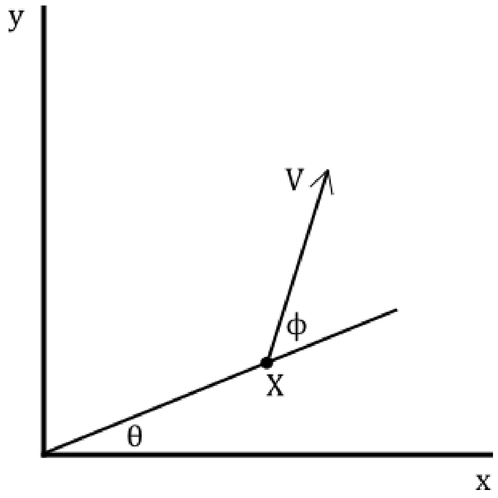

Figure 3: A vector V based at point X in $2D$ plane polar coordinates.

Since the plane is two-dimensional, there are two ways write a parallel-transported version of V, which shall be denoted with a tilde symbol, namely $\tilde{V}$. Note V and $\tilde{V}$ are identical in magnitude and direction. Begin by finding expressions for V^r, V^θ in terms of the coordinates r, θ by requiring that the contraction $V^\mu V_\mu$ be to equal the square of the magnitude $|V|^2 = |\tilde{V}|^2$. Using the $2D$ plane polar metric (171), we have

$$|V|^2 = g_{\alpha\beta}V^\alpha V^\beta = g^{\mu\nu}V_\mu V_\nu \,,$$

implying

$$V^r = |V|\cos\phi \qquad\qquad V^\theta = \frac{1}{r}\,|V|\sin\phi$$
$$V_r = |V|\cos\phi \qquad\qquad V_\theta = r\,|V|\sin\phi\,.$$

Let the basepoint X undergo the small translation $r \to r + \Delta r$, effectively sliding V along the radial line away from the origin. The (contravariant) components of the parallel-transported vector $\tilde{V}$ are

$$\tilde{V}^r = \left|\tilde{V}\right|\cos\phi \qquad\qquad \tilde{V}^\theta = \frac{1}{r+\Delta r}\left|\tilde{V}\right|\sin\phi\,.$$

Enforcing the fact that V and $\tilde{V}$ are equivalent, use $V^\theta = \tilde{V}^\theta$ to attain

$$\tilde{V}^\theta = \frac{r}{r+\Delta r}V^\theta \approx V^\theta - V^\theta\frac{\Delta r}{r}\,, \tag{199}$$

where the translations Δr are vey small compared to r.

Repeat the same exercise for small arc-shaped 'translations' of the basepoint X in the variable θ such that $\theta \to \theta + \Delta\theta$. This causes the angle ϕ to decrease by $\phi \to \phi - \Delta\theta$. In this case, we write

$$\tilde{V}^r = \left|\tilde{V}\right| \cos\left(\phi - \Delta\theta\right) \qquad\qquad \tilde{V}^\theta = \frac{1}{r}\left|\tilde{V}\right| \sin\left(\phi - \Delta\theta\right) \, .$$

Enforcing the fact that V and $\tilde{V}$ are equivalent, use $V^r = \tilde{V}^r$ to attain

$$\tilde{V}^r = |V| \left(\cos\phi - \Delta\theta \sin\phi\right) \approx V^r - V^\theta r \Delta\theta \, , \tag{200}$$

where $\Delta\theta$ is much less than ϕ.

Equations (199) and (200) hint at something more general: it appears that the parallel-transported vector $\tilde{V}^\mu$ under general coordinate variations dq^μ looks like

$$\tilde{V}^\mu \approx V^\mu + V^\lambda \left(?\right)^\mu_{\nu\lambda} dq^\nu \, ,$$

but what do we use for the unknown obect? Since it's structure is suspisciously Γ-like, so let's boldly try

$$\tilde{V}^\mu \approx V^\mu + V^\lambda \Gamma^\mu_{\nu\lambda} dq^\nu \, . \tag{201}$$

Looking up the nonzero $\Gamma^\mu_{\nu\lambda}$ for $2D$ plane polar coordinates, namely $\Gamma^\theta_{r\theta} = 1/r$ and $\Gamma^r_{\theta\theta} = -r$, equation (201) spits out equations (199) and (200) again! You can repeat this exercise for any of the coordinate systems explored so far - it always works.

Assuming equation (201) should work in general, we gain a new insight into what the Christoffel symbols really are: for small parallel transport in coordinates x^μ, each $\Gamma^\mu_{\nu\gamma}$ tells us how much a vector's components blend togaether, despite that the vector V and its new version $\tilde{V}$ are the same. Non-Cartesian coordinates are responsible for this.

Going a step ahead, observe that equation (201) has the appearance of the calculus-one formula for the derivative of V^μ. If V^μ is the four-velocity, equal to equal to $dq^\mu/d\tau$, then dividing through by τ yields

$$0 = \frac{d^2 q^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda} \frac{dq^\nu}{d\tau} \frac{dq^\lambda}{d\tau} \, ,$$

formally known as the geodesic equation.

Problem 109: Consider $2D$ plane polar coordinates with the metric given by $dS^2 = dr^2 + r^2 d\theta^2$. A vector vield V^μ has unit length and points in the x direction everywhere. (i) Find its components V^r and V^θ. (ii) Show that the vector field of part (i) satisfies $D_\mu V^\nu = 0$ in polar coordinates. Answer:

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} V^\nu \qquad\qquad V^r = \cos\theta \qquad\qquad V^\theta = \frac{-\sin\theta}{r}$$

$$D_r V^r = D_r V^\theta = D_\theta V^r = D_\theta V^\theta = 0$$

9.10 Geodesic Equation

Derivation by Parallel Transport

Let us address parallel transport for a completely general situation. Consider some four-vector V^μ that varies with generalized coordinates q^μ on a manifold, and the coordinates

q^μ are written in terms of the parameter λ. (Note that time *is* included in this analysis.) Parallel-transporting the vector V^μ in any direction amounts to requiring that its derivative with respect to λ must be zero:

$$\frac{dV^\mu}{d\lambda} = \frac{dq^\alpha}{d\lambda} D_\alpha V^\mu \left(q^\mu\left(\lambda\right)\right) = 0 \tag{202}$$

For scalar fields, equation (202) boils down to $\partial\phi/\partial\lambda = dq^\alpha/d\lambda \cdot \partial_\alpha \phi$, which is the product of the tangent vector and the gradient.

Now supposing V^μ is the velocity four-vector, we know V^μ is defined as the τ-derivative of the generalized position vector q^μ, where τ is the proper time. (The specific relationship between λ and τ is not assumed just yet.) Using the chain rule, equation (202) can be manipulated to say

$$-\frac{dq^\alpha}{d\lambda}\frac{dq^\mu}{d\lambda}\frac{\partial}{\partial q^\alpha}\left(\frac{d\lambda}{d\tau}\right) = \left(\frac{d\lambda}{d\tau}\right)\left(\frac{d^2 q^\alpha}{d\lambda^2} + \Gamma^\mu_{\alpha\gamma}\frac{dq^\gamma}{d\lambda}\frac{dq^\alpha}{d\lambda}\right) . \tag{203}$$

Equation (203) is known as the *geodesic equation* with a *non-affine parameterization*, and is admittedly ugly. The trouble starts with λ, which is so far unspecified. Suppose though that λ's relationship to τ is *affine*, specified by

$$\lambda = b_0 \tau + a_0 . \tag{204}$$

In such a case, the left side of (203) vanishes, allowing us to write the extremely handy *affinely-parameterized geodesic equation*:

$$\frac{d^2 q^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho}\frac{dq^\nu}{d\tau}\frac{dq^\rho}{d\tau} = 0 \tag{205}$$

Right away, we know equation (205) is useful because (i) the geodesic equation is solved by worldlines of constant four-velocity, and (ii) a worldline with constant four-velocity corresponds to inertial motion. In other words, the geodesic equation is the curved-space generalization of $F = ma$, but for $F = 0$.

I'll warn that most authors on this subject, if not all of them, will write dx in place of dq in the geodesic equation. The use of dx-notation is fine if you don't mind dx^μ having dimensions other than length.

Parallel Transport on Two-Sphere

Problem 110: Consider a circle C on the two-sphere (the surface of an ordinary sphere) of constant angle θ_0. At ϕ_0 the vector V^μ has components $V^\theta = 0$, $V^\phi = 1$. Compute the components of V^μ as a function of ϕ if it is parallel-transported around C. Answer:

$$0 = \frac{dV^\theta}{d\phi} - V^\phi \sin\theta\cos\theta \qquad\qquad 0 = \frac{dV^\phi}{d\phi} + V^\theta\frac{\cos\theta}{\sin\theta}$$

$$V^\theta = \sin\theta_0 \cdot \cos\left[\phi\cdot\cos\theta_0\right] \qquad\qquad V^\phi = \cos\left[\phi\cdot\cos\theta_0\right]$$

Derivation by Variational Method

The geodesic equation may also be derived by variational methods, starting with the spacetime interval $dS^2 = g_{\mu\nu}dq^\mu dq^\nu$. Replacing dS^2 by $-c^2 d\tau^2$, we have, after solving for $d\tau$,

$$d\tau = \frac{1}{c}\sqrt{-g_{\mu\nu}dq^\mu dq^\nu} .$$

Defining the action

$$S = \int_{\tau_i}^{\tau_f} c \cdot d\tau \ ,$$

and then minimizing both sides with respect to small variations in τ will yield a kind of Euler-Lagrange equation that happens to be the geodesic equation. Supposing the coordinates q^μ are expressed in terms of τ, the action is

$$S = \int_{\tau_i}^{\tau_f} d\tau \sqrt{-g_{\mu\nu} \frac{dq^\mu}{d\tau} \frac{dq^\nu}{d\tau}} \ .$$

Introducing small variations in the generalized coordinates, let q^μ vary as $q^\mu \to q^\mu + \delta q^\mu$, making S go to

$$S + \delta S = \int_{\tau_i}^{\tau_f} d\tau \sqrt{-g_{\mu\nu} \frac{d}{d\tau} \left(q^\mu + \delta q^\mu\right) \frac{d}{d\tau} \left(q^\nu + \delta q^\nu\right)} \ . \tag{206}$$

The metric $g_{\mu\nu}$ is actually a function of the coordinates q^μ. To first order expansion, $g_{\mu\nu}$ behaves has

$$g_{\mu\nu} \approx g_{\mu\nu} + \delta q^\alpha \frac{\partial}{\partial q^\alpha} g_{\mu\nu} \ .$$

Boiling the algebra inside the integral, the square root term reduces to

$$\approx \sqrt{c^2 - \delta q^\alpha \frac{\partial g_{\mu\nu}}{\partial q^\alpha} \frac{dq^\mu}{d\tau} \frac{dq^\nu}{d\tau} - g_{\mu\nu} \frac{d\delta q^\mu}{d\tau} \frac{dq^\nu}{d\tau} - g_{\mu\nu} \frac{dq^\mu}{d\tau} \frac{d\delta q^\nu}{d\tau}} \ ,$$

where any terms involving the multiplication of two small δ-terms are ignored, and the relation

$$-c^2 = g_{\mu\nu} \frac{dq^\mu}{d\tau} \frac{dq^\nu}{d\tau}$$

has been used. Furthermore, c^2 is a number far greater than any other term inside the square root, so we may approxmiate

$$\sqrt{c^2 - A} \approx c - \frac{A}{2c} \ .$$

As the dust settles, notice a factor of S drops off of each side of equation (206), leaving

$$\delta S = \frac{1}{2c} \int_{\tau_i}^{\tau_f} d\tau \left(\delta q^\alpha \frac{\partial g_{\mu\nu}}{\partial q^\alpha} \frac{dq^\mu}{d\tau} \frac{dq^\nu}{d\tau} + g_{\mu\nu} \frac{d\delta q^\mu}{d\tau} \frac{dq^\nu}{d\tau} + g_{\mu\nu} \frac{dq^\mu}{d\tau} \frac{d\delta q^\nu}{d\tau} \right) \ . \tag{207}$$

To handle the integral above, integrate the second and third terms by parts. For the third term (and similarly for the second term), you might set

$$u = g_{\mu\nu} \frac{dq^\mu}{d\tau} \qquad du = \frac{dg_{\mu\nu}}{d\tau} \frac{dq^\mu}{d\tau} + g_{\mu\nu} \frac{d^2 q^\mu}{d\tau^2}$$

$$dv = \frac{d\delta q^\nu}{d\tau} d\tau \qquad v = \delta q^\nu$$

such that

$$\int u dv = uv\big| - \int v du \ .$$

Invoking the assumption that the variation in τ vanishes at the points τ_i and τ_f, the boundary term completely vanishes. The third term on the right side of equation (207) becomes, after using the chain rule and shuffling some indices,

$$\delta S_3 = -\frac{1}{2c}\int_{\tau_i}^{\tau_f} d\tau\,(\delta q^\alpha)\left(\frac{dg_{\alpha\mu}}{dq^\nu}\frac{dq^\nu}{d\tau}\frac{dq^\mu}{d\tau} + g_{\mu\alpha}\frac{d^2q^\mu}{d\tau^2}\right).$$

Repeating the process for the second term in (207) shoud give

$$\delta S_2 = -\frac{1}{2c}\int_{\tau_i}^{\tau_f} d\tau\,(\delta q^\alpha)\left(\frac{dg_{\nu\alpha}}{dq^\mu}\frac{dq^\mu}{d\tau}\frac{dq^\nu}{d\tau} + g_{\mu\alpha}\frac{d^2q^\mu}{d\tau^2}\right).$$

The updated version of equation (207) reads

$$\delta S = \frac{1}{c}\int_{\tau_i}^{\tau_f} d\tau\,(\delta q^\alpha)\left[g_{\mu\alpha}\frac{d^2q^\mu}{d\tau^2} + \frac{1}{2}\left(\frac{\partial g_{\mu\nu}}{\partial q^\alpha} - \frac{\partial g_{\nu\alpha}}{\partial q^\mu} - \frac{\partial g_{\alpha\mu}}{\partial q^\nu}\right)\left(\frac{dq^\mu}{d\tau}\frac{dq^\nu}{d\tau}\right)\right].$$

Next make the crucial observation that δS should be vanishingly small for a geodesic curve, making the left side of the above equation tend to zero. This is only possible if the term in square brackets is zero:

$$0 = g_{\mu\alpha}\frac{d^2q^\mu}{d\tau^2} + \frac{1}{2}\left(\frac{\partial g_{\mu\nu}}{\partial q^\alpha} - \frac{\partial g_{\nu\alpha}}{\partial q^\mu} - \frac{\partial g_{\alpha\mu}}{\partial q^\nu}\right)\left(\frac{dq^\mu}{d\tau}\frac{dq^\nu}{d\tau}\right)$$

Contracting with $g^{\rho\alpha}$ gives

$$0 = \frac{d^2q^\rho}{d\tau^2} + \frac{1}{2}g^{\rho\alpha}\left(\frac{\partial g_{\mu\nu}}{\partial q^\alpha} - \frac{\partial g_{\nu\alpha}}{\partial q^\mu} - \frac{\partial g_{\alpha\mu}}{\partial q^\nu}\right)\left(\frac{dq^\mu}{d\tau}\frac{dq^\nu}{d\tau}\right).$$

Having equation (194) already memorized, notice there is a Christoffel symbol $\Gamma^\rho_{\mu\nu}$ in the above. Condensing the result down one more time:

$$0 = \frac{d^2q^\rho}{d\tau^2} + \Gamma^\rho_{\mu\nu}\frac{dq^\mu}{d\tau}\frac{dq^\nu}{d\tau}$$

You will recognize this as the geodesic equation (205).

Christoffel Symbols by Variational Method

Problem 111: Determine the nonzero Christoffel symbols for FRW space by the variational method. Do so by minimizing

$$I = \int g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}d\tau$$

with the metric given in equation (181). (Vary parameters t and x separately.) Compare your result to the geodesic equation to reproduce the Christoffel symbols (195).

Geodesics in Minkowski Space

A trivial (nonzero) solution to the geodesic equation is straight lines for Minkowski space. The Minkowski metric $\eta_{\mu\nu}$ has no derivative, and the geodesic equation becomes a second order differential equation,

$$\frac{d^2x^\mu}{d\tau^2} = 0\,.$$

Integrating twice with respect to τ we find the equations of a straight line:

$$x^{\mu}(\tau) = x_0^{\mu} + U_0^{\mu}\tau$$

Of course, the constant U_0^{μ} must be normalized according to $-c^2 = \eta_{\mu\nu}U^{\mu}U^{\nu}$, namely

$$U_0^0 = \pm\sqrt{c^2 + \vec{U}_0 \cdot \vec{U}_0} \, .$$

Geodesics in the Plane

Geodesics in the Cartesian plane be described by

$$x(\lambda) = a\lambda + x_0 \qquad\qquad y(\lambda) = b\lambda + y_0 \, ,$$

where λ is a dimensionless parameter. In $2D$ plane polar coordinates, variables (x, y) relate to (r, θ) by

$$x = r\cos\theta \qquad\qquad y = r\sin\theta \, .$$

Thus, straight lines in polar coordinates look like

$$r(\lambda) = \sqrt{(a\lambda + x_0)^2 + (b\lambda + y_0)^2} \qquad\qquad \theta(\lambda) = \tan^{-1}\left(\frac{b\lambda + y_0}{a\lambda + x_0}\right) \, .$$

Using what we know about $\Gamma^{\mu}_{\nu\rho}$ in $2D$ plane polar coordinates, the geodesic equation (205) yields two bits of information:

$$\frac{d^2r}{d\lambda^2} - r\left(\frac{d\theta}{d\lambda}\right)^2 = 0 \qquad\qquad \frac{d^2\theta}{d\lambda^2} + \frac{2}{r}\frac{dr}{d\lambda}\frac{d\theta}{d\lambda} = 0$$

Problem 112: Prove that $r(\lambda)$ and $\theta(\lambda)$ given above are solutions to the geodesic equation for $2D$ plane polar coordinates.

Geodesics on Two-Sphere

Problem 113: Write down the geodesic equation on the two-sphere and prove that all lines of constant ϕ (latitude) are geodesics, while the lines of constant θ are only geodesic on the equator. Answer:

$$\frac{d^2\theta}{d\tau^2} - \sin\theta\cos\theta\left(\frac{d\phi}{d\tau}\right)^2 = 0 \qquad\qquad \frac{d^2\phi}{d\tau^2} + \cot\theta\frac{d\theta}{d\tau}\frac{d\phi}{d\tau} = 0$$

Geodesics for Massless Particles

A massless particle according to general relativity can *only* be a photon, and travels at speed c. The velocity four-vector for a photon satisfies the normalization condition

$$0 = g_{\nu\rho}dq^{\nu}dq^{\rho} \, . \tag{208}$$

Meanwhile, the geodesic equation says

$$0 = \frac{d^2q^{\mu}}{d\lambda^2} + \Gamma^{\mu}_{\nu\rho}\frac{dq^{\nu}}{d\lambda}\frac{dq^{\rho}}{d\lambda} \, .$$

Notice that the parameter λ is used instead of the proper time τ, because for photons the proper time is identically zero. The solution $q^\mu(\lambda)$ is called a *null geodesic*.

We may as well let λ equal the time t as observed by *some* reference frame, so long as $\tau \neq t$. In such a case, we write

$$0 = g_{\mu\nu} \frac{dq^\mu}{dt} \frac{dq^\rho}{dt} \, .$$

Demonstrating on a general $1+1D$ example, the above equation reads

$$0 = g_{00} + 2g_{01} \left(\frac{dq}{dt} \right) + g_{11} \left(\frac{dq}{dt} \right)^2 \, ,$$

which is an equation of motion easily solved for $q(t)$. For the trivial case of Minkowski space, we immediately see $dq/dt = c$, as expected.

Let us repeat the same exercise for $1+1D$ Rindler coordinates, where the metric (158) reads

$$dS^2 = -c^2 d\tau^2 \left(1 + \frac{\xi \tilde{a}}{c^2} \right)^2 + d\xi^2 \, ,$$

and remember $dS^2 = 0$ for photons, and τ is the proper time at the 'floor' of the accelerated frame (call it a rocket) where $\xi = 0$. At the floor of the rocket, the above reduces to

$$\frac{d\xi}{d\tau} = c \, ,$$

which is no surprise. Even an accelerated observer should see photons moving at speed c.

It's trickier to consider observers above the floor of the rocket, where we have

$$\frac{d\xi}{d\tau} = c \left(1 + \frac{\xi \tilde{a}}{c^2} \right) \, ,$$

which is deceptive because the right side isn't plain c. Recall that the local time differential in Rindler space is given by equation (161), which says

$$d\tau_\xi = d\tau \left(1 + \frac{\xi \tilde{a}}{c^2} \right) \, .$$

What we're really interested in is the quantity $d\xi / d\tau_\xi$, which comes to

$$\frac{d\xi}{d\tau_\xi} = \frac{d\xi}{d\tau} \frac{d\tau}{d\tau_\xi} = c \, ,$$

as expected.

9.11 Riemann Curvature Tensor

So far, no construction has been proposed that can be used to discern whether or not a given metric $g_{\mu\nu}$ corresponds to curved space or flat space. The problem is that when you are embedded in one space or another, it appears locally flat. Some kind of procedure must be done to try to 'detect' curved space. For instance, walking consistently westward you might discover that the Earth is at least cylindrical.

As mentioned previously, the interior degrees of a triangle don't necessarily sum to 180° in curved space. Imagine laying out a great triangle on the Earth as follows: (i)

From the north pole, draw an arc to any point on the equator. (ii) From that point, trace out a distance on the equator comparable in length to the first arc. (iii) Connect the end of the second arc to the north pole, completing the 'triangle'. The sum of each interior angles will surpass 180°.

Now modify the triangle-drawing exercise with a new idea. As you trace out the triangle on the two-sphere, suppose you 'carry along' a vector $\vec{V}$ without allowing rotations, meaning that $\vec{V}$ is parallel-transported along the triangular circuit. Here's the kicker: when $\vec{V}$ returns to its starting point, its final oreintation is *different* than its initial orientation. To an observer embedded on the surface, where all seems flat, it would appear very strange that $\vec{V}$ should change its orientation. This is how one could detect a curved world, as $\vec{V}$ will *never* rotate or rotate on a true plane.

To quantify the notion of *curvature*, consider a manifold mapped by coordinates q^μ. What we will do is parallel-transport a vector on a small closed path, and check for rotations in $\vec{V}$. Begin with the geodesic equation (203), and expand it in a way that resembles equation (201), namely

$$dV^\rho + \Gamma^\rho_{\mu\nu} dq^\mu V^\nu = 0 \, .$$

Take V^μ as the vector that is parallel-transported along the circuit segments dq^ν on the manifold. For a circuit, we will use a small closed parallelogram-shaped path on the manifold with vertices located at q^0, q^1, q^2, q^3. Denoting the displacement from q^j to q^k as a^μ_{jk}, the above equation tells us

$$dV^\rho\left(a_{01}\right) = V^\rho\left(q^1\right) - V^\rho\left(q^0\right) = -\Gamma^\rho_{\mu\nu}\left(0\right) a^\mu_{01} V^\nu\left(q^0\right) \, .$$

The second path involves going from q^1 to q^2, and we write

$$dV^\rho\left(a_{12}\right) = V^\rho\left(q^2\right) - V^\rho\left(q^1\right) = -\Gamma^\rho_{\mu\nu}\left(1\right) a^\mu_{12} V^\nu\left(q^1\right) \, .$$

Remember the Christofel symbols Γ are themselves functions of the coordinates q^μ at points q^j. Thus in certain instances, Christoffel symbols are reported as $\Gamma^\mu_{\alpha\beta}(j)$. In terms of $V^\rho(q^0)$, the quantity $V^\rho(q^2)$ is equal to

$$V^\rho\left(q^2\right) = V^\rho\left(q^0\right) - V^\nu\left(q^0\right)\left(\Gamma^\rho_{\mu\nu}\left(0\right) a^\mu_{01} - \Gamma^\rho_{\mu\nu}\left(1\right) a^\mu_{12} + \Gamma^\rho_{\mu\epsilon}\Gamma^\epsilon_{\alpha\nu} a^\mu_{01} a^\alpha_{12}\right) \, .$$

Next, the third path involves going from q^2 to q^3, meaning

$$dV^\rho\left(a_{23}\right) = V^\rho\left(q^3\right) - V^\rho\left(q^2\right) = -\Gamma^\rho_{\mu\nu}\left(2\right) a^\mu_{23} V^\nu\left(q^2\right) \, .$$

The final path goes from q^3 back to q^0, so we have

$$dV^\rho\left(a_{30}\right) = V^\rho\left(q^0\right) - V^\rho\left(q^3\right) = -\Gamma^\rho_{\mu\nu}\left(3\right) a^\mu_{30} V^\nu\left(q^3\right) \, .$$

Some things need clarification as you trudge through this algebraic mess becaue everything must be expressed in terms of $V^\mu(q^0)$. First, the quantites a^μ_{jk} are small, so when we see the product of three of these, the term is ignored. Secondly, the derivative of a Christoffel symbol can be written as

$$a^\alpha_{jk}\partial_\alpha \Gamma^\mu_{\nu\rho} \approx \Gamma^\mu_{\nu\rho}\left(k\right) - \Gamma^\mu_{\nu\rho}\left(j\right) \, . \tag{209}$$

Thirdly, the paths a_{01} and a_{23} are equal and opposite, and in the same way, so too are a_{12} and a_{30}.

As the last step, construct the total change in the vector V^μ, which is

$$\delta V^\mu = V^\mu\left(q^4\right) - V^\mu\left(q^0\right) \ .$$

If everything was done very carefully, you should arrive at

$$\delta V^\rho = \left[\partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda \nu\sigma - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}\right] a^\mu_{01} a^\nu_{12} V^\sigma\left(q^0\right) \ .$$

Stripping away the 'fluff' terms we write down

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma} \ , \tag{210}$$

an object called the *Riemann curvature tensor*. As promised, $R^\rho_{\sigma\mu\nu}$ is what we use to determine whether or not a surface has curvature. All components of the Riemann curvature tensor are zero for flat space.

Properties of the Riemann Curvature Tensor

As a four-index object, the Riemann tensor in general has $4^4 = 256$ individual components. Due to certain symmetries in the Riemann tensor though, the number of independent components is reduced significantly. Here we write down four important proerties of $R_{\rho\sigma\mu\nu}$ (check them yourself):

- $R_{\alpha\beta\gamma\rho}$ is antisymmetric in the last two indices:

$$R_{\alpha\beta\gamma\rho} = -R_{\alpha\beta\rho\gamma} \tag{211}$$

- $R_{\alpha\beta\gamma\rho}$ is antisymmetric in the first two indices:

$$R_{\alpha\beta\gamma\rho} = -R_{\beta\alpha\gamma\rho} \tag{212}$$

- $R_{\alpha\beta\gamma\rho}$ is symmetric when exchanging pairs of indices as follows:

$$R_{\alpha\beta\gamma\rho} = R_{\gamma\rho\alpha\beta} \tag{213}$$

- The Riemann tensor obeys:

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\gamma\beta\delta} = 0 \tag{214}$$

Using equations (211) - (214), it can be shown that the number of independent components in the Riemann tensor becomes

$$N = \frac{1}{12} D^2\left(D^2 - 1\right) \ ,$$

where D is the number of dimensions. For example, $D = 1$ can't possibly involve curvature, and we correspondingly find $R_{1111} = 0$. In two dimensions, there is only one independent component of the Riemann tensor, namely $R_{0101} \neq 0$.

Bianchi Identity

The covariant derivative operator D_α applies straightforwardly to the Riemann tensor $R_{\beta\gamma\delta\epsilon}$. It's relatively easy to to show that

$$D_\alpha R_{\beta\gamma\delta\epsilon} + D_\beta R_{\gamma\alpha\delta\epsilon} + D_\gamma R_{\alpha\beta\delta\epsilon} = 0 \ , \tag{215}$$

known as the *Bianchi identiy.*

Analogy to Electromagnetism

In the case of electroagnetism, information on electric and magnetic fields is stored in the electromagnetic field strength tensor $F_{\mu\nu}$. Various operations on $F_{\mu\nu}$ yield the equations of motion, energy density, and so on for electromagnetism. For spacetime geometry, the same can be said for the Riemann tensor. To draw the analogy, recall the 'Maxwell' equation (118) that involves the four-curl of the electromagnetic field strength tensor, as in

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} = 0\,,$$

which looks very much like the Bianchi identiy (215).

Also in electromagnetism we apply guage transformations to A^μ, for instance $A^\mu \to A^\mu + \partial^\mu \alpha$, where α is a scalar field. In a sense, this is analogous to coordinate changes that bring the metric from $g_{\mu\nu}$ to $g'_{\mu\nu}$.

9.12 Ricci Tensor and Ricci Scalar

Recall that the Riemann curvature tensor carries information about the curvature of a given manifold. Being a four-index object however, the Riemann curvature tensor can be a rather unruly object. It would be handy to be able to 'access' curvature information on a manifold by using an object simpler than $R^\alpha_{\beta\delta\gamma}$. The tool for this job is called the *Ricci curvature tensor*, which is a particular contraction of the Riemann curvature tensor as follows:

$$R_{\mu\nu} = R^\alpha_{\mu\alpha\nu} \tag{216}$$

The Ricci curvature tensor $R_{\mu\nu}$ describes the growth rate of volume elements on a manifold. Being a special case of $R^\alpha_{\beta\delta\gamma}$, it would be repetitive to spell out all of the identities relating to $R_{\mu\nu}$.

Going a step further, the trace of $R_{\mu\nu}$, namely

$$R = g^{\alpha\beta} R_{\alpha\beta} \tag{217}$$

is known as the *Ricci scalar curvature*. If you know something about curvature already, R is not genrally equivalent to the mean curvature or Gaussian curvature for a surface. In two dimensions though, it turns out that the Ricci scalar curvature is twice the Gaussian curvature.

9.13 Einstein Tensor

A few manipulations of the Bianchi identity yields the ever-important *Einstein tenosr*. To motivate the problem, suppose we seek some contraction and/or combination of factors involving the Ricci tensor that satisfies

$$D_\alpha G^\alpha_\beta = 0 \qquad\qquad D_\gamma G^{\gamma\rho} = 0\,, \tag{218}$$

i.e., zero divergence in $(1,1)$ and $(2,0)$ forms.

Begin with equation (215) and multiply thtrough by $g^{\delta\rho}$, giving

$$D_\alpha R^\rho_{\beta\gamma\epsilon} + D_\beta R^\rho_{\gamma\alpha\epsilon} + D_\gamma R^\rho_{\alpha\beta\epsilon} = 0\,.$$

Multiply through by $g^{\alpha\epsilon}$, and we get

$$D_\alpha R^{\gamma\alpha}_{\beta\gamma} + D_\beta R^{\rho\alpha}_{\gamma\alpha} + D_\gamma R^{\rho\alpha}_{\alpha\beta} = 0\,.$$

Next, exploit the antisymmetry properties of R and also let $\rho = \gamma$:

$$-D_\alpha R^{\gamma\alpha}_{\gamma\beta} + D_\beta R^{\gamma\alpha}_{\gamma\alpha} - D_\gamma R^{\gamma\alpha}_{\beta\alpha} = 0$$

Observe that the first and third terms are equal (after a change in indices), and the second term reduces to the divergence of the Ricci scalar.

So far then,

$$D_\alpha R^\alpha_\beta - \frac{1}{2} D_\beta R = 0 \,,$$

which is a total divergence:

$$D_\alpha \left(R^\alpha_\beta - \frac{1}{2} g^\alpha_\beta R \right) = 0$$

Raising the β index and plucking out the term in parenthesis, the Einstein tensor is evidently

$$G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \,. \tag{219}$$

9.14 Geodesic Deviation

Consider a family of geodesics characterized by $q^\mu(\tau, s)$, where τ is the proper time and s is an arc length parameter. Define the (small) *deviation vector* S^μ such that

$$q^\mu \to q^\mu + S^\mu \,,$$

where S^μ is perpendicular to the (tangent) four-velocity $U^\mu = dq^\mu / d\tau$ at any point q^μ. The geodesic equation (205) under deviations $q^\mu \to q^\mu + S^\mu$ appears as

$$\frac{d^2 (q^\mu + S^\mu)}{d\tau^2} + \Gamma^\mu_{\alpha\beta} (q^\mu + S^\mu) \frac{d (q^\alpha + S^\alpha)}{d\tau} \frac{d \left(q^\beta + S^\beta\right)}{d\tau} = 0 \,.$$

Subtracting the undeviated ($S^\mu = 0$) geodesic equation from the above yields a handy identity. To pick it out, note that the Christoffel symbol $\Gamma^\mu_{\alpha\beta}$ evaluated over a small interval can be written in terms of its derivative, as first done in equation (209). The above boils down to

$$\frac{d^2 S^\mu}{d\tau^2} + 2\Gamma^\mu_{\nu\rho} U^\nu \frac{dS^\rho}{d\tau} + U^\lambda U^\gamma S^\epsilon \frac{\partial}{\partial q^\epsilon} \Gamma^\mu_{\lambda\gamma} = 0 \,. \tag{220}$$

Next define the *relative velocity* V^μ and *relative acceleration* A^μ, respectively, as

$$V^\mu = \frac{DS^\mu}{d\tau} = U^\alpha D_\alpha S^\mu \qquad\qquad A^\mu = \frac{DV^\mu}{d\tau} = U^\alpha D_\alpha V^\mu \,.$$

Inserting V^μ into the A^μ-equation, we find $A^\mu = U^\alpha D_\alpha \left(U^\gamma D_\gamma S^\mu\right)$, expanding to

$$A^\mu = U^\alpha \left(D_\alpha U^\gamma\right) D_\gamma S^\mu + U^\alpha U^\gamma D_\alpha D_\gamma S^\mu \,.$$

On the right side, the first term is identically zero, as

$$U^\alpha D_\alpha U^\gamma = U^\alpha \left(\partial_\alpha U^\gamma + \Gamma^\gamma_{\alpha\beta} U^\beta \right) = \frac{dU^\gamma}{d\tau} + \Gamma^\gamma_{\alpha\beta} U^\alpha U^\beta = 0 \,, \tag{221}$$

which is just a restatement of the geodesic equation. Expanding out the term $U^\alpha U^\gamma D_\alpha D_\gamma S^\mu$, you should have

$$A^\mu = U^\alpha U^\gamma D_\alpha D_\gamma S^\mu = U^\alpha U^\gamma S^\nu \left(-\partial_\nu \Gamma^\mu_{\alpha\gamma} + D_\alpha \Gamma^\mu_{\gamma\nu} + \Gamma^\mu_{\gamma\beta} \Gamma^\beta_{\alpha\nu} \right)$$

after replacing the second derivative terms by equation (220). Next write down the relation

$$D_\alpha \Gamma^\mu_{\gamma\nu} = \partial_\alpha \Gamma^\mu_{\gamma\nu} + \Gamma^\mu_{\alpha\beta}\Gamma^\beta_{\gamma\nu} - \Gamma^\beta_{\alpha\gamma}\Gamma^\mu_{\beta\nu} - \Gamma^\beta_{\alpha\nu}\Gamma^\mu_{\beta\gamma} \, .$$

Evidently, A^μ is equal to

$$A^\mu = U^\alpha U^\gamma S^\nu \left[-\partial_\nu \Gamma^\mu_{\alpha\gamma} + \partial_\alpha \Gamma^\mu_{\gamma\nu} + \Gamma^\mu_{\alpha\beta}\Gamma^\beta_{\gamma\nu} - \Gamma^\beta_{\alpha\gamma}\Gamma^\mu_{\beta\nu} \right] \, ,$$

where you should recognize the term in square brackets as the Riemann curvature tensor $R^\mu_{\gamma\alpha\nu}$.

The above analysis may be regarded as an alternative derivation of the Riemann curvature tensor (210). The star result however is the 'simple' formula for A^μ, known as the *geodesic deviation equation*:

$$\frac{d^2 S^\mu}{d\tau^2} = A^\mu = R^\mu_{\gamma\alpha\nu} U^\gamma U^\alpha S^\nu \tag{222}$$

Equation (222) is not to be interpreted as an equation of motion (force on a particle, for instance). Instead, the geodesic deviation equation describes how spacetime curvature influences adjacent geodesics. Don't forget that the proper time $d\tau$ is in a sense equivalent to the arc dS that goes along geosesics, which is handy for a purely geometric interpretation of general relativity.

9.15 Killing Vectors

There is an important difference between an invariant and a constant of motion (in spacetime). That is, the the norm of the velocity four-vector is equal to $-c^2$ for all coordinate systems, but this does not imply that the three-velocity is a constant. Similarly, the norm of the four-momentum is always $-m^2c^2$, but this is not to be confused with a statement about momentum conservation.

The real statements about conservation of momentum, energy, or otherwise come from Noether 's theorem, embodied in equation (89), allowing one to find a conserved current j^μ associated with *some* change to the Lagrangian density $\mathcal{L}$. For instance, simple translations in spacetime, namely $x^\mu \to x^\mu - a^\mu$, leads to the stress-energy tensor, as equation (90) reminds us:

$$\partial_\mu j^\mu = 0 \qquad\qquad j^\mu = a_\nu T^{\mu\nu} \qquad\qquad \partial_\mu T^{\mu\nu} = 0$$

The translation vector a_ν is actually called a *Killing vector* (named after Wilmelm Killing). It turns out that a conserved quantity tensor on a manifold is always associated with some killing vector. Notice that our conversation on classical field theory took place in flat space, so the ∂_μ operator in the above must be replaced by the covariant derivative, D_μ. Writing out $D_\mu j^\mu$, we have

$$0 = D_\mu j^\mu = T^{\mu\nu} D_\mu a_\nu + a_\nu D_\mu T^{\mu\nu} \, ,$$

where the second term is zero by construction. Due to symmetry in $T^{\mu\nu}$, we have

$$0 = D_\mu j^\mu = \frac{1}{2} T^{\mu\nu} \left(D_\mu a_\nu + D_\nu a_\mu \right) \, .$$

The term in parenthesis must be equal to zero. In the general case, T is replaced by any conserved tensor, and the Killing vector a_ν shall be generally written X_ν. Any Killing vector must obey *Killing's equation*,

$$D_\mu X_\nu + D_\nu X_\mu = 0\,. \tag{223}$$

Killing's equation tells us that a_ν, or more generally generally X_ν, is a vector along which global translations (or rotations, etc.) of the manifold leave the metric invariant.

Metric Isometry (something to fix here WFB)

Consider the coordinate transformation $x^\mu \to \tilde{x}^\mu$ such that

$$x^\mu = \tilde{x}^\mu + \lambda^\mu\,,$$

such that λ is small. Using the transformation law

$$\tilde{g}_{\alpha\beta} = \frac{\partial x^\mu}{\partial \tilde{x}^\alpha}\frac{\partial x^\nu}{\partial \tilde{x}^\beta} g_{\mu\nu}\,,$$

it's possible to show that

$$\tilde{g}_{\alpha\beta} = g_{\alpha\beta} + D_\alpha \lambda_\beta + D_\beta \lambda_\alpha\,,$$

and if λ_μ is porportional to a Killing vector K_α, then $\lambda_\mu = \lambda K_\alpha$.

Conservation of Four-Momentum

A particle that moves along a geodesic in spacetime has constant four-momentum, meaning P^μ is parallel-transported along the geodesic $x^\mu(\tau)$. The geodesic equation (205) predicts the past and future motion of the particle in terms of the proper time τ.

Consider a small vector K_μ along which we translate the entire spacetime. If the contraction between P^μ and K_μ has zero τ-derivative along the geodesic, we know (i) that K_μ is a Killing vector, and (ii) $P^\mu K_\mu$ is a constant of motion. Calculating $d_\tau(P^\mu K_\mu)$, we start with

$$\frac{d}{d\tau}\left(P^\mu K_\mu\right) = mU^\alpha D_\alpha \left(U^\mu K^\mu\right) = mK_\mu U^\alpha D_\alpha U^\mu + mU^\mu U^\alpha D_\alpha K_\mu\,,$$

where U^μ is the four-velocity of the particle. The $U^\alpha D_\alpha U^\mu$ term is identically zero by the geodesic equation. Due to the symmetry in $U^\mu U^\alpha$, we write

$$\frac{d}{d\tau}\left(P^\mu K_\mu\right) = \frac{m}{2} U^\mu U^\alpha \left(D_\alpha K_\mu + D_\mu K_\alpha\right)\,,$$

which evaluates to zero by Killing's equation (223). The corresponding constant of motion shall be denoted $\tilde{E}$, as in

$$\tilde{E} = P^\mu K_\mu\,,$$

where the tilde ($\tilde{}$) symbol is a reminder that the dimensionality is not Joules. Choosing a different Killing vector associated with rotations, denoted R_μ, we gain another constant of motion related to angular momentum:

$$\tilde{L} = P^\mu R_\mu$$

10 Newtonian Gravity

10.1 Inventing the Theory

Newton's law of gravitation did not begin with a statement about the force on an apple, nor did he simply blurt out the universal law - he was actually interested in describing celestial motions in terms of central attractions. Realizing the scope of such a problem, Newton had to simultaneously invent the tools for the job: calculus plus the laws of motion. We shall follow a modernized narrative of Newton's discovery of the universal law.

Drawing on the work of Kepler, Newton knew that orbiting bodies follow elliptical trajectories given by *Kepler's first law of ellipses*, published in 1609, which we take to be

$$r\left(\theta\right) = \frac{1}{A - B\cos\theta} \; .$$

The form for $r(\theta)$ contains the information that the origin is the center of attraction, called the *focus*, yet the constants A and B were not understood by Kepler. Also in 1609, Kepler popularized his *second law*, stating that *a line drawn between the sun and the planet sweeps out equal areas in equal times as the planet orbits the sun.*

Newton invented what we know today as the angular momentum vector, namely $\vec{L} = \vec{r} \times \vec{p}$, where $\vec{r}$ is the displacement vector from the origin to the orbiting body, and $\vec{p}$ is the linear momentum, the product of the mass and the velocity vector $\vec{v}$. Also available to Newton was his own second law, stating that the force vector $\vec{F}$ is equal to the mass times the acceleration vector. Taking the time derivative of $\vec{L}$, we find

$$\frac{d}{dt}\vec{L} = m\left(\vec{v} \times \vec{v}\right) + m\left(\vec{r} \times \frac{d^2\vec{r}}{dt^2}\right) \; .$$

The $\vec{v} \times \vec{v}$ term is zero by the rules of cross products, and the second term may be rewritten to say

$$\frac{d}{dt}\vec{L} = \vec{r} \times \vec{F} = 0 \; ,$$

which evaluates to zero because the force vector and displacement vector are along the same line. It follows that $\vec{L}$ is a constant of motion, and it's magnitide is given by

$$L = mr^2 \frac{d\theta}{dt} \; .$$

That is, a body in orbit has constant angular momentum. Looking at our formula for L, rearrange to get

$$r^2 \frac{d\theta}{dt} = \frac{L}{m} \; .$$

The left side describes (twice) the rate of area sweep in polar coordinates, and the right side is constant - this is precisely the mathematical statement of Kepler's second law.

Newton's advantage was clearly the use of derivatives, so in that spirit we calculate the slope at a point on any ellipse, along with the second derivative:

$$\frac{d}{dt} r\left(\theta\right) = -Br^2 \sin\theta \frac{d\theta}{dt}$$

$$\frac{d^2}{dt^2} r\left(\theta\right) = -B\sin\theta \frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) - Br^2\cos\theta\left(\frac{d\theta}{dt}\right)^2 = \frac{\left(L/m\right)^2}{r^2}\left(\frac{1}{r} - A\right)$$

Meanwhile, it's easily shown that the radial acceleration of a point in polar coordinates obeys

$$a_r = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 .$$

Substituting what's known about $d\theta/dt$ and d^2r/dt^2, arrive at

$$a_r = -\frac{L^2 A}{m^2}\frac{1}{r^2} \propto -\frac{1}{r^2} .$$

It appears that the acceleration of an orbiting body is (i) purely radial, (ii) attractive, and (iii) has magnitude proportional to the inverse square of the position vector. The product of the mass m and the gravitational acceleration of the orbiting body is proportional to the gravitational force on the body:

$$F \propto -\frac{m}{r^2}$$

By Newton's own third law, the above statement must also be true for the attracting object (the sun), which has its own mass M. This means we're up to $F \propto -Mm/r^2$. Newton polished his result with a proportionality constant G (which was not known until the Cavendish experiment took place in 1798). *Newton's universal law of gravitation* reads:

$$\vec{F} = -\frac{GMm}{r^2}\hat{r} \tag{224}$$

(The constant L relates to G, M, etc. as $L = m\sqrt{GM}/\sqrt{A}$.)

In 1618, Kepler published his *third law*, stating that *the square of the period of a planet is directly proportional to the cube of the semi-major axis of its orbit.* It must have taken Kepler hundreds of hours of to notice this. Using the laws of motion and gravitation, Newton was able to reproduce Kepler's third law, and also determine the correct proportionality constants.

The ellipse, described by $r(\theta)$ above, contains two important length scales. The largest diameter across the ellipse is defined as two times the *semimajor axis*, and similarly, the *semiminor axis* is half of the shortest diameter of the ellipse. Denote the semi- major and minor axes as a and b, respectively. It's straightforward to show that a and b relate to A and B by:

$$a = \frac{(1/A)}{1-(B/A)^2} \qquad\qquad b = \frac{(1/A)}{\sqrt{1-(B/A)^2}}$$

To calculate the period T of a planetary orbit, integrate both sides of $r^2 d\theta/dt = L/m$ with respect to time over one complete orbit, giving

$$2\pi ab = \frac{L}{m}T .$$

On the left, we get twice the area of the ellipse. Rewriting L and b in terms of the semimajor axis, we find the precise statement of Kepler's third law:

$$T = 2\pi\frac{a^{3/2}}{\sqrt{GM}}$$

10.2 Analogy to Electromagnetism

The gravitational force between two objects has a form identical to that of the electrostatic force between point charges. Both force laws are proportional to the product of two 'charges', and inversely proportional to the inverse square of the distance separating them. Of course, electromagnetism consists of more than just the force between charges; the subject is enriched by notions such as the electric potential, electric flux, and charge density - all tied up neatly in vector calculus. In this section we'll frame Newtonian gravity in a similar way.

Begin by writing the gravitational force bewteen two objects, denoted m and M, as follows:

$$\vec{F} = -\frac{GMm}{r^2}\hat{r}$$

Dividing by the 'test mass' m gives the gravitational field

$$\vec{g} = -\frac{GM}{r^2}\hat{r} \; ,$$

whose magnitude on Earth's surface is the familiar $9.8\,m/s^2$.

Enclosing the mass M inside a Gaussian surface $\mathcal{S}$, the gravitational flux Φ_G may be written

$$\Phi_G = \int_{\mathcal{S}} \vec{g}\cdot\hat{N}\,dA = -GM\int_{\mathcal{S}} \frac{\cos\theta}{r^2}\,dA \; ,$$

where θ is the angle between the field $\vec{g}$ and the surface normal $\hat{N}$. The integral covers the entire solid angle enclosing the object M, and without loss of generality we may let the Gaussian surface be spherical. We find that

$$\Phi_G = -GM\int_{-1}^{1}\cos\theta\,d\left(\cos\theta\right)\int_0^{2\pi} d\phi = -4\pi GM \; ,$$

stating that the gravitational flux through a surface is proportional to the total mass enclosed. This is the gravitational equivalent to Gauss's law from electromagnetism.

Applying the divergence theorem to the Φ_G equation, we write

$$\int_{\mathcal{S}} \vec{g}\cdot\hat{N}\,dA = \int_{\tau}\vec{\nabla}\cdot\vec{g}\,dV = -4\pi G\int_{\tau}\rho\,dV \; .$$

The total mass M has been written as the volume integral of the mass density $\rho(\vec{r})$ over the region τ enclosed by $\mathcal{S}$. Stripping away the volume integrals we have

$$\vec{\nabla}\cdot\vec{g} = -4\pi G\rho \; ,$$

the differential version of Gauss's law for graviation.

Next, we notice that the gravitational field $\vec{g}$ is curl-free, meaning $\vec{\nabla}\times\vec{g}$ is identically zero. It follows that the field $\vec{g}$ can be written as the gradient of a scalar gravitational potential, denoted $V(\vec{r})$ such that

$$\vec{g} = -\vec{\nabla}V(\vec{r}) \; .$$

Substituting $\vec{g}$ into the differential Gauss law above, we arrive at Poisson's equation

$$\nabla^2 V(\vec{r}) = 4\pi G\rho \; , \tag{225}$$

which is directly analogous to (110) from electrodynamics. The gravitational potential is the line integral of the gravitational field:

$$V(\vec{r}) = \int_{\infty}^{\vec{r}} \vec{g}(\vec{r}') \cdot (-\hat{r})\, dr' = -\frac{GM}{r}\,, \tag{226}$$

where $V(\infty)$ is taken as zero.

10.3 Shortcomings of Newtonian Gravity

Most of the shortcomings of Newtonian physics, particularly at high energies, are corrected by the special theory of relativity. All seems well until we try to account for gravity. First off, gravity becomes nonlinear: a gravitational field, which carries energy, must carry additional mass due to the energy-mass equivalence $E = mc^2$. The additional mass has a gravitational field of its own, which contributes more energy, and so on.

Another problem with Newtonian gravity is that Newton's equation seems to work independent of time, implying that gravitational signals may be sent instantaneously over arbitrary distances. No other force in physics behaves this way. The well-accepted *standard model* predicts accurately that the other forces of physics are mediated by particles that move at speeds no greater than c. Although the standard model makes no mention of gravity, it should make sense that we adopt a speed limit of c for gravitational signals.

Yet another motivation for improving Newton's theory of gravitation was the precession of planet Mercury's orbit around the sun. It turns out that *every* planet's orbit isn't quite right if we strictly follow Newton, but only Mercury is close enough to the sun to test the Newtonian theory for strong gravitational fields. Indeed, when astronomers looked carefully at Mercury, it's orbit did not precess as Newton predicted.

11 Gravity in Curved Space

Our studies of geometry in curved space have availed a number of lessons that apply directly to physics. Equations of motion for light and matter are actually trajectories in spacetime called geodesics. Further, there is some connection between the metric $g_{\mu\nu}$ and the gravitational potential $V(\vec{x})$, implying that Christoffel symbols are somehow related to the notion of force and acceleration.

11.1 Weak-Field Gravity

For the moment we examine the weak-field nonrelativistic limit, which to say we examine the physics of a pressureless lump of dust at rest in a weak gravitational field. Let's not suppose that the metric $g_{\mu\nu}$ be flat, but instead picks up a small time-independent perturbative term

$${}^{(\text{weak})}g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{227}$$

such that $h_{\mu\nu} \ll \eta_{\mu\nu}$.

<u>Problem 114</u>: The up-index version of ${}^{(\text{weak})}g_{\mu\nu}$ is not $\eta^{\mu\nu} + h^{\mu\nu}$. Expand $g^{\mu\nu} = \eta^{\mu\nu} + H^{\mu\nu}$ to several orders in $h_{\mu\nu}$ by two different methods: (i) Derive and use recursively the

relation $H^{\mu\rho} = -\eta^{\mu\nu}g^{\sigma\rho}h_{\nu\sigma}$. (ii) Use a λ-power expansion $g_{\mu\nu} = \eta_{\mu\nu} + \lambda h_{\mu\nu}$, $H^{\mu\nu} = \lambda H^{\mu\nu}_{(1)} + \lambda^2 H^{\mu\nu}_{(2)} + \cdots$. Answer:

$$^{(\text{weak})}g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} + h^{\mu}_{\sigma}h^{\sigma\nu} - h^{\mu}_{\sigma}h^{\sigma}_{\beta}h^{\beta\nu} + \cdots \tag{228}$$

In the weak-field nonrelativistic limit, the position vector $q^\mu = (ct, \vec{x})$ has corresponding four-velocity $U^\mu = (c\gamma, \partial_\tau \vec{x}) \approx (c, 0, 0, 0)$. Meanwhile, the geodesic equation (205) tells us

$$\frac{d^2 q^\mu}{d\tau} + \Gamma^{\mu}_{00}\frac{dq^0}{d\tau}\frac{dq^0}{d\tau} = 0\,.$$

The Christoffel symbols Γ^{μ}_{00} are calculated from equation (194), giving

$$\Gamma^{\mu}_{00} = -\frac{1}{2}g^{\mu\alpha}\partial_\alpha g_{00} \approx -\frac{1}{2}\eta^{\mu\alpha}\partial_\alpha h_{00}\,,$$

where only first-order terms in $h_{\mu\nu}$ are retained. Combining the two equations, arrive at

$$\frac{d^2 q^\mu}{d\tau^2} = \frac{c^2}{2}\eta^{\mu\alpha}\partial_\alpha h_{00}\,.$$

The $\mu = 0$ case (time component) tells us $t = t_0 + \tau$. Using this fact while taking the $\mu = j$ case (spatial components), arrive at

$$\frac{\partial^2 x^j}{\partial t^2} = \frac{c^2}{2}\partial_j h_{00}\,,$$

or in vector form,

$$\frac{d^2\vec{x}}{dt^2} = \frac{c^2}{2}\vec{\nabla}h_{00}\,. \tag{229}$$

In the above, the left side is none other than the local acceleration due to gravity, also known as $\vec{g}$. The right side not only brings the speed of light into the picture, but also puts a restriction on the metric perturbation $h_{\mu\nu}$.

More generally, equation (229) may be written

$$h_{00} = -\frac{2}{c^2}V(\vec{x})\,, \tag{230}$$

where $V(\vec{x})$ is the gravitational potential. In other words, we find that the gravitational potential may be expressed as some part of the metric!

Problem 115: In the weak-field nonrelativistic limit, show that the 00-component of the Ricci tensor obeys:

$$R_{00} \approx -\frac{1}{2}\partial^\alpha\partial_\alpha h_{00} = -\frac{1}{2}\nabla^2 h_{00} \tag{231}$$

Problem 116: The 00-component of the uniform-acceleration Rindler space metric (158) reads

$$g_{00} = -\left(1 + \frac{\xi\tilde{a}}{c^2}\right)^2\,.$$

In the small-$\xi\tilde{a}$ limit, check that the corresponding gravitational potential is equal to $\xi\tilde{a}$.

11.2 Einstein's Field Equation

The motivation for our next steps comes from the analogy between electromagnetism and gravitation, with the common entity being Poisson's equation, applying to both regimes in the generic form

$$\nabla^2 V = k\rho \,,$$

where V is some kind of potential, ρ is a density, and k is constant.

The natural question is, what is a fully generalized version of Poisson's equation? On the right side, we need to replace the density ρ with something that has all the right properties: energy and momentum conservation, covariance, and so on. To our great luck, the stress energy tensor $T^{\mu\nu}$ fits this description perfectly!

Meanwhile, on the left side, we need an object that behaves something like $T^{\mu\nu}$: a $(0, 2)$ symmetric tensor. In addition, the left side of the equation demands a term that is two spatial derivatives of a potential or metric, which loosely speaking, is equivalent to one spatial derivative of the Christoffel symbols, landing us in the ballpark of the Riemann and Ricci tensors. Looking back at our previous work, it appears that the Einstein tensor (219) fits the bill. In its derivation, we specifically sought a combination of factors that satisfy $D_\mu G^{\mu\nu} = 0$, which must be true for anything proportional to $T^{\mu\nu}$.

We therefore write

$$G^{\mu\nu} \propto T^{\mu\nu} \,.$$

It tells us the Einstein tensor, which is merely a slice of information about the curvature of space, is proportional to the stress-energy tensor, which maps the actual distribution of matter and charges in space. Einstein is asking us to believe that the distribution of matter determines the curvature of space, and vice versa: the curvature of space determines how the matter contained will behave. In general relativity, the phenomenon of gravity is completely explained in terms of curved space instead of traditional forces.

To determine the exact form of Einstein's field equation, the proportionality constant κ in $G^{\mu\nu} = \kappa T^{\mu\nu}$ must be determined. We're free to pin down the κ any way we like, so take the easiest route possible: the weak-field nonrelativistic limit. Begin with the low-index version of the Einstein tensor (219), and contract with $g^{\mu\nu}$, giving

$$g^{\mu\nu} G_{\mu\nu} = g^{\mu\nu} \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) = \kappa g^{\mu\nu} T_{\mu\nu} \,.$$

Now I point out three things: (i) $g^{\mu\nu} R_{\mu\nu}$ is the Ricci scalar R; (ii) $g^{\mu\nu} g_{\mu\nu}$ counts the dimensions of the space at hand, in this case, 4; and (iii) $g^{\mu\nu} T_{\mu\nu}$ is the trace of $T_{\mu\nu}$, called T. From this, we find

$$-R = \kappa T \,,$$

such that the field equation now reads

$$R_{\mu\nu} + \frac{1}{2} g_{\mu\nu} \kappa T = \kappa T_{\mu\nu} \,.$$

Recalling our discussion on fluids, the stress-energy tensor for a perfect fluid at rest is trivial. All pressure, momentum, and stress components are zero; only the 00-component remains, and is given by $T_{00} = c^2 \rho$, where ρ is the mass density. The trace of $T_{\mu\nu}$ is equal to $-c^2 \rho$ due to the minus sign on the 00-component of the flat metric. It follows that

$$R_{00} = \frac{1}{2} c^2 \kappa \rho \,,$$

and the job is essentially done. We already know R_{00} from equation (231) in terms of a perturbed metric. Eliminating the perturbation h_{00} by equation (230), we find

$$\frac{c^4\kappa}{2}\rho = \nabla^2 V\left(\vec{x}\right) \ ,$$

which is just Poisson's equation for graviation. For this to be true, κ must equal $8\pi G/c^4$, where G is Newton's gravitation constant. Finally then, the central equation of general relativity - the mighty Einstein field equation - reads:

$$G^{\mu\nu} = \frac{8\pi G}{c^4}T^{\mu\nu} \tag{232}$$