General Relativity Study Notes

These notes attempt to summarize the overall formalism, assuming that you have already seen things like the Einstein summation convention and the occurence of raised and lowered indices.

The basic assumption of general relativity is that times and distances are described by a pseudo-Riemannian metric $g_{\mu\nu}$. And that, in the abscence of forces other than gravity, particles follow geodesics of the geometry. The metric itself is determined by the distribution of energy (and pressure) through Einstein's equation.

The fact that space-time is curved means that we have to be very careful that the quantities we write down are independent of the co-ordinate system that we are working in.

I. PARTICLE TRAJECTORIES, LENGTHS AND TIMES: SPECIAL RELATIVITY REVIEW

Here are some facts that work in all geometries. We will consider a particle path in space-time described by the paramter λ $x^\mu(\lambda)$ where μ runs over the four space-time indices. We will use Greek indices to indicate the four-spacetime indices starting from $\mu = 0$ for the time index and Latin indices to indicate the spatial indices only.

Notice that we will use a variety of co-ordinates x^μ in this course t, r, θ, ϕ for example in spherical polar co-ordinates.

We will sometimes use the bold-face notation for a four-vector and an arrow to indicate a three-vector $\mathbf{x} = (t, x, y, z) = (t, \vec{x})$. We will often change the basis and co-ordinates that we use to describe vectors so we can think of a general four-vector as a sum over basis vectors like this (for $\mathbf{x}$ and for a general four-vector $\mathbf{a}$)

$$\mathbf{x} = t\mathbf{e}_0 + x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 \quad (1)$$
$$\mathbf{a} = a^0\mathbf{e}_0 + a^1\mathbf{e}_1 + a^2\mathbf{e}_2 + a^3\mathbf{e}_3 \quad (2)$$
$$(3)$$

Now if we specialize to general relativity for a moment. If we have two points $\mathbf{x}_1$ and $\mathbf{x_2}$ such that $\Delta\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_1$then the invariant interval that describes the "distance" between them is

$$\Delta s^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \quad (4)$$
$$\equiv \eta_{\mu\nu} x^\mu x^\nu \equiv \mathbf{\Delta x} \cdot \mathbf{\Delta x} \quad (5)$$

Here I have defined both the metric for flat spacetime $\eta = \mathrm{diag}(-c^2, 1, 1, 1)$ and the dot-product notation for four-vectors. Shortly I will choose to use units such that $c = 1$ since this greatly simplifies the formulae that we will use.

Now Δs^2 may be positive, indicating a space-like interval, zero indicating a null interval or negative indicating a time-like interval.

There is an infinitesimal version of this metric that we will often resort to in curved spacetime

$$ds^2 = -(cdt)^2 + (dx)^2 + (dy)^2 + (dz)^2. \quad (6)$$

Recall that the time experienced by an observer or clock that moves at constant velocity over a time-like interval $\mathbf{\Delta x}$ is given by

$$\Delta\tau^2 = -\Delta s^2/c^2 = -\eta_{\mu\nu}\Delta x^m \Delta x^\nu. \quad (7)$$

If the clock is not moving at constant velocity we need to add up the contributions of many such intervals that are short enough in time that the velocity is approximately constant. This will give us an integral for the proper time

$$\tau_{ab} = \int_A^B d\tau \equiv \int_A^B d\sqrt{dt^2 - (dx^2 + dy^2 + dz^2)/c^2} \quad (8)$$
$$= \int_A^B dt \sqrt{1 - \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]/c^2} \quad (9)$$
$$= \int_A^B dt \sqrt{1 - |\vec{v}|^2(t)/c^2}. \quad (10)$$

There is a differential shorthand for this equation

$$d\tau = dt\sqrt{1 - |\vec{v}|^2(t)/c^2} \tag{11}$$

This expresses the fact that in special relativity moving clocks run slow.

Notice that this formula works perfectly well for accelerating observers. It makes it particularly easy to see why in the "twin paradox" it is the twin who goes off on a spaceship who returns younger. We could consider the two twins A and B one of whom B who stays on Earth on an inertial trajectory. Since the interval we are interested in is time-like there is a reference frame in which B is at rest and experiences a time $\tau_B = t_2 - t_1$. In this frame A is not at rest and experiences a time

$$\tau_A = \int_{t_1}^{t_2} dt\sqrt{1 - |\vec{v}|^2(t)/c^2} \leq \tau_B \tag{12}$$

where the less than or equal to is because $\sqrt{1 - |\vec{v}|^2(t)/c^2}$ is always less than one. The assymetry between the two twins is that while B is always in an inertial frame, A is not.

As an aside the geodesic paths in flat space time are straight and B's path is one. We can see from this that the geodesic is the longest path between two time-like separated points. This is in marked contrast to the case in flat Euclidean space.

Recall that we can change from one reference frame to another by a Lorentz transformation (into a frame moving at velocity v along the x-axis)

$$\begin{aligned} ct' &= \gamma(t - vx/c) = \cosh\theta(ct) - \sinh\theta x && (13)\\ x' &= \gamma(x - vt) = -\sinh\theta(ct) + \cosh\theta x && (14)\\ y' &= y \quad z' = z && (15) \end{aligned}$$

Here we have used the usual gamma factor

$$\gamma = (1 - \vec{v}^2(t)/c^2)^{-1/2} \tag{16}$$

and have written the transformation in terms of the alternative variable, the rapidity θ which is define such that

$$v = c\tanh\theta. \tag{17}$$

The expressions in terms of the rapidity highlight the sense in which a Lorentz boost is in some way comparable to a rotation.

The Lorentz transformation is a linear transformation and in the general case we will use the matrix Λ to indicate a Lorentz transformation. For a general four-vector we have

$$\mathbf{a}' = \Lambda\mathbf{a} \tag{18}$$

or in terms of components

$$a'^{\mu} = \Lambda^{\mu}{}_{\nu}a^{\nu}. \tag{19}$$

You should check that under this transformation the invariant interval Δs is unchanged, so that $\Delta s' = \Delta s$. Things are defined so that this is the case. Because of this it is also easy to show that for any two four-vectors we have

$$\mathbf{a}' \cdot \mathbf{b}' = \mathbf{a} \cdot \mathbf{b} \tag{20}$$

so that dot products are independent of the reference frame we use to evaluate them in. This property will also carry over to curved spacetimes.

Just a quick note to remind you of the *relativity of simultaneity.* If two events are space-like separated then there exist reference frames where either one occurs first. The two events cannot be causally connected.

Lets consider as an example a uniformly accelerated observer who has the following path in spacetime, parameterized by σ

$$t(\sigma) = a^{-1}\sinh\sigma \quad x(\sigma) = a^{-1}\cosh\sigma. \tag{21}$$

Notice that the path obeys $x^2 - t^2 = a^{-2}$ so it is a hyperbola.

We can work out how proper time evolves along this path by using

$$d\tau^2 = dt^2 - dx^2. \tag{22}$$

(dy and dz are zero because the y and z co-ordinates are not changing with time.) Which we should interpret as

$$\left(\frac{d\tau}{d\sigma}\right)^2 = \left(\frac{dt}{d\sigma}\right)^2 - \left(\frac{dx}{d\sigma}\right)^2 \tag{23}$$

$$= (a^{-1}\cosh\sigma)^2 - (a^{-1}\sinh\sigma)^2 \tag{24}$$

$$= a^{-2}. \tag{25}$$

So we can conclude that $\tau = \sigma/a$ if $\tau = 0$ when $\sigma = 0$. This allows us to write the path using τ the proper time as the parameter rather than σ. We find

$$t(\tau) = a^{-1}\sinh a\tau \quad x(\tau) = a^{-1}\cosh a\tau. \tag{26}$$

There is a general message here. For a particle trajectory it is always possible to parameterize the path in terms of the proper time τ. The calculation proceeds by exactly the steps we have used here.

The four velocity $\mathbf{u}$ for a particle trajectory given by $x^\mu(\tau)$ is defined to be

$$\mathbf{u} = \frac{d\mathbf{u}}{d\tau} \tag{27}$$

or in terms of components

$$u^\mu = \frac{dx^\mu}{d\tau}. \tag{28}$$

We can likewise define the three velocity

$$\vec{v} = \frac{d\vec{x}}{dt}. \tag{29}$$

Notice that this is a derivative with respect to co-ordinate time, not proper time.

Notice as well that the definition of the proper time 72 is just

$$\frac{dt}{d\tau} = \gamma \tag{30}$$

Likewise the chain rule gives us

$$\frac{dx}{d\tau} = \frac{dt}{d\tau}\frac{dx}{dt} \tag{31}$$

so we can conclude that

$$\mathbf{u} = (\gamma, \gamma\vec{v}). \tag{32}$$

Finally we can calculate the length of the velocity four-vector

$$\mathbf{u}\cdot\mathbf{u} = \eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = -\left(\frac{d\tau}{d\tau}\right)^2 = -1 \tag{33}$$

The fact that the norm squared of $\mathbf{u}$ is negative means that we call $\mathbf{u}$ a *timelike* four-vector. This normalisation condition will be used many times over in the what follows.

Looking at our example of an accelerated observer we find $u^0 = dt/d\tau = \cosh a\tau$ and $u^1 = dx/d\tau = \sinh a\tau$ and we find that this does indeed satisfy $\mathbf{u}\cdot\mathbf{u} = -1$. From this we also find the velocity as a function of proper time $v = dx/dt = \tanh a\tau$. Notice how at negative proper times this approaches the speed of light in the negative x direction and at positive proper times it approaches the speed of light in the positive x direction.

A. Forces and four-momenta

We come now to the rules of mechanics. Newton's first law is that particle velocities do not change unless a force is exerted. In special relativity, this should indicate that the four-velocity does not change with proper time unless a force is exerted. That is in the absence of a force

$$\frac{d\mathbf{u}}{d\tau} = 0. \tag{34}$$

Newton's second law should relate accelerations to forces. We will use m to indicate the (rest) mass of a particle. Then we can define the acceleration four-vector

$$\mathbf{a} \equiv \frac{d\mathbf{u}}{d\tau} \tag{35}$$

and we might guess that there is a four vector force $\mathbf{f}$ such that

$$\mathbf{f} = m\mathbf{a} \tag{36}$$

Now on the face of it this seems different from classical mechanics in that there appear to be four equations of motion rather than three. This is not really the case since the length of the four-velocity does not change so we have

$$m\frac{d\mathbf{u}\cdot\mathbf{u}}{d\tau} = 0 = 2m\mathbf{u}\cdot\frac{d\mathbf{u}}{d\tau} \tag{37}$$

from which we can conclude that the conditions

$$\mathbf{f}\cdot\mathbf{u} = 0 = \mathbf{a}\cdot\mathbf{u} \tag{38}$$

and there are in fact only three independent equations just as in classical equations.

Lets look at our example. We find $a^0 = du^0/d\tau = a\sinh a\tau$ and $a^1 = du^1/d\tau = a\cosh a\tau$. This gives us $\mathbf{a}\cdot\mathbf{a} = a^2$ so the parameter a can indeed be regarded as an acceleration. Likewise we find $\mathbf{a}\cdot\mathbf{u} = 0$ as required. So this trajectory is one which is undergoing a constant acceleration along the x-axis.

We can define the four-momentum as follows

$$\mathbf{p} = m\mathbf{u}. \tag{39}$$

The time component of the four-momentum we will take to be the energy $p^0 = E = m\gamma$ while the spatial components we will write as $\vec{p} = m\gamma\vec{v}$. In terms of the four-momentum Newton's second law may be written

$$\mathbf{f} = \frac{d\mathbf{p}}{d\tau} \tag{40}$$

exactly as we expect. Notice that

$$\mathbf{p}\cdot\mathbf{p} = m^2\mathbf{u}\cdot\mathbf{u} = -m^2 \tag{41}$$

this in turn means that we find

$$E = \sqrt{m^2 + \vec{p}^2} \tag{42}$$

This should be a familiar relativistic formula.

The time component of Newton's second law has the interpretation that the change in energy is equal to the work done. To see this we can define

$$\vec{F} \equiv \frac{d\vec{p}}{dt} \tag{43}$$

in terms of which we have $\vec{f} = \gamma\vec{F}$ (since $dt/d\tau = \gamma$). If we look at the condition on the force four-vector we find

$$0 = \mathbf{f}\cdot\mathbf{u} = -\gamma f^0 + \gamma\vec{f}\cdot\vec{v} \tag{44}$$

from which we can conclude

$$f^0 = \vec{f}\cdot\vec{v} = \gamma\vec{F}\cdot\vec{v} \tag{45}$$

so we get for the time component of Newton's second law

$$\frac{dE}{dt} = \frac{df^0}{d\tau}/\gamma = \vec{F} \cdot \vec{v}. \tag{46}$$

Finally lets just check the interpretation that p^0 is the energy. If $v \ll 1$ we can use a binomial expansion for γ and we find

$$p^0 = m + m\vec{v}^2/2 + ... \tag{47}$$

So indeed E looks like rest mass energy plus kinetic energy in the limit of low velocity.

B. Light

Consider the path of a photon, travelling at the speed of light along the x-axis

$$x = t \tag{48}$$

Notice that $v = 1$, this is what determines that this is a light path.

We can express this in terms of a parameter λ as

$$\mathbf{x}(\lambda) = \lambda \mathbf{u} \tag{49}$$

where $\mathbf{u} = (1, 1, 0, 0)$. Notice that

$$\mathbf{u} = \frac{d\mathbf{x}}{d\lambda} \tag{50}$$

and $d\mathbf{u}/d\lambda = 0$ so that the light path has no "acceleration". Such a parameterization of the light path is known as an *affine parameterization.*

We find that

$$\mathbf{u} \cdot \mathbf{u} = 0. \tag{51}$$

Four-vectors of zero length like this are called *null* vectors. The four-vector acceleration of light, or any massless particle, is a null vector.

We can take a slightly more classical view of a trajectory of light by considering a classical plane wave. We'll generalize to allow the light to travel in any direction. A plane wave of light has electric field proportional to

$$e^{-i(\omega t - \vec{k} \cdot \vec{x} + \delta)} = e^{-i(k_\mu x^\mu + \delta)} \tag{52}$$

where we have defined the wave four vector

$$\mathbf{k} = \begin{pmatrix} \omega \\ \vec{k} \end{pmatrix} \tag{53}$$

Since we can write the phase of the wave as a scalar product it is Lorentz invariant. Recall that $\omega = c|\vec{k}|$ so we have (in units where $c = 1$)

$$\mathbf{k} \cdot \mathbf{k} = k_\mu k^\mu = \omega^2 - |k|^2 = \omega^2 - \omega^2 = 0. \tag{54}$$

So the wave four vector is a null vector.

The wavefronts move at the speed of light in the direction given by $\vec{k}$ and are described by the condition

$$k_\mu x^\mu + \delta = 2\pi n \tag{55}$$

Lets choose a point $x^\mu_s = (t_s, x_s, y_s, z_s)$ on the wavefront with $n = 1$ (that is $k_\mu x^\mu_s + \delta = 2\pi$). Consider the trajectory

$$x^\mu(\lambda) = \begin{pmatrix} \omega\lambda + t_s \\ k_1\lambda + x_s \\ k_2\lambda + y_s \\ k_3\lambda + z_s \end{pmatrix}. \tag{56}$$

Notice firstly that this point travels along with the wavefront since

$$k_\mu x^\mu(\lambda) + \delta = \lambda k_\mu k^\mu + k_\mu x_s^\mu + \delta = 2\pi. \tag{57}$$

Notice secondly that this point on the wave front moves at the speed of light in the direction of $\vec{k}$ since

$$v^i = \frac{dx^i}{d\lambda}\frac{d\lambda}{dt} = k^i/\omega \tag{58}$$

which is a vector $\vec{v} = \vec{k}/\omega$ having length one and pointing in the correct direction.

Notice finally that

$$\frac{dx^\mu}{d\lambda} = k^\mu \tag{59}$$

so the four-velocity for a plane wave of light, with this parameterization λ is just the wave four vector k^μ.

So we see that at least for this choice of λ we can interpret the four-velocity $d\mathbf{x}/d\lambda$ of light as the wave four vector $\mathbf{k}$.

Now we can determine the momentum four vector for light most easily by recalling the quantum mechanical expressions for the energy and momentum of a photon

$$E = \hbar\omega \tag{60}$$

$$\vec{p} = \hbar\vec{k} \tag{61}$$

or in four vector notation

$$p^\mu = \hbar k^\mu. \tag{62}$$

Notice that we may parameterize our wavefront trajectory in terms of a new variable λ' for which $\lambda = \hbar\lambda'$ (notice λ' needs to have appropriate units to compensate for this)

$$x^\mu(\lambda') = \begin{pmatrix} \hbar\omega\lambda' + t_s \\ \hbar k_1\lambda' + x_s \\ \hbar k_2\lambda' + y_s \\ \hbar k_3\lambda' + z_s \end{pmatrix}. \tag{63}$$

This trajectory is identical to the previous one (which you can check by investigating dx/dt for example) but now we have

$$p^\mu = \frac{dx^\mu}{d\lambda'}. \tag{64}$$

This means that we can choose to parameterize a null trajectory in such a way that its four velocity is equal to its four momentum. This is often a convenient parameterization, although the previous one is also natural.

In general the bottom line for light is that

$$\mathbf{u} \propto \mathbf{k} \propto \mathbf{p}. \tag{65}$$

All three of these vectors are null and we can arrange of $\mathbf{u}$ with either of the other two by an appropriate choice of λ.

II. OBSERVED QUANTITIES

Suppose an observer at rest in our reference frame measures the energy of a particle with four momentum $\mathbf{p}$. The measured energy is of course $p^0 = E$.

This formula is not terrible satisfying since it requires picking out a component of some four-vector and these components will have different values in different reference frames. However in principle we should be able to do this calculation in any reference frame with the outcome being the same. This suggests that it should be possible to express the measured energy in terms of some dot product that would take the same value in all reference frames.

You are used to doing these calculations by boosting the co-ordinates to the reference frame in which the observer is at rest and taking the appropriate component of $\mathbf{p}$. This will not be convenient in general relativity where spaces are curved and so we are motivated to find such an expression in terms of dot products.

The way we do this is to write the four-vector velocity of the observer. Since the observer is at rest we know that $dx^i/d\tau = 0$ so we only need to find $dt/d\tau$ but this is set by the normalisation of $\mathbf{u}$ to be one. Thus we have

$$\mathbf{u}_{\rm obs} = (1, 0, 0, 0) \tag{66}$$

and we can write the measured energy as

$$E = -\mathbf{p} \cdot \mathbf{u}_{\rm obs}. \tag{67}$$

(Check this by using the definition of the dot product.) Since the expression on the right hand side is a dot product it is true in any reference frame.

In special relativity this formula is overkill but it is still true in general relativity and there it is indispensible.

Lets test out this formula by considering the energy observed by an observer moving with velocity $\vec{v}$ of a particle of mass m at rest. We have $\mathbf{p} = (m, 0, 0, 0)$ and $\mathbf{u}_{\rm obs} = (\gamma, \gamma\vec{v})$ and thus we find the observered energy to be γm. But this is exactly what we expect. The measured mass is greater than the rest mass because in the reference frame of the observer the particle is moving with a velocity v.

We can use the same trick to find measured frequencies, which will allow us to calculate redshifts in general relativity. The same argument that we used for massive particles allows us to find the measured frequency of a light pulse in terms of the four-velocity of the observer

$$\omega = -\mathbf{k} \cdot \mathbf{u}_{\rm obs}. \tag{68}$$

Lets consider the frequency of light from a star at rest in our co-ordinate system as observed by our accelerated observer. Imagine that the starlight is propagating along the x-axis. We find $\mathbf{k} = (\omega^*, \omega^*, 0, 0)$. Then the observed frequency is

$$\omega(\tau) = -\mathbf{k} \cdot \mathbf{u}_{\rm obs} = k^0 u^0 - k^1 u^1 = \omega^*(\cosh a\tau - \sinh a\tau) = \omega^* \exp(-a\tau). \tag{69}$$

So the starlight is blue-shifted at early times and red-shifted at late times. Indeed the star is only visible for a time of order $1/a$ since at other times the Doppler shift will result in it being outside the observerable frequencies.

We might also be interested in the magnitude of the velocity of a particle with four-vector velocity $\mathbf{v}$ as measured by an observer with four-velocity $\mathbf{u}_{\rm obs}$. Again in the reference frame in which the observer is at rest we can pick this out from the time component of the four-velocity $|v| = \sqrt{1 - \gamma^{-2}}$. And of course $\gamma = -\mathbf{v} \cdot \mathbf{u}_{\rm obs}$. Thus we can write the measured velocity as

$$|v| = \sqrt{1 - \mathbf{v} \cdot \mathbf{u}_{\rm obs}^{-2}}. \tag{70}$$

This formula involves only constants and a scalar product and as a result is independent of co-ordinates.

III. CURVED SPACE VERSION OF THIS

This just recapitulates what I have said earlier.

We may define the four-vector $dx^\mu/d\lambda$. If

$$g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} < 0 \tag{71}$$

then the particle trajectory is *timelike.* All actual physical trajectories of massive particles are timelike.

For timelike trajectories it makes sense to ask how much time a clock moving along with the particle would advance during the motion. This quantity is usually known as the “proper time of the particle” and the assumption is that this is given us by the metric

$$\tau = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}. \tag{72}$$

On the other hand if

$$g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} > 0 \tag{73}$$

then the particle trajectory is *spacelike*. In this case the particle trajectory can be assigned a physical length. This length is likewise given by the metric

$$s = \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}. \tag{74}$$

Finally if

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0 \tag{75}$$

the particle trajectory is null. Light and other massless particles travel null trajectories.

IV. FOUR-VECTOR VELOCITIES AND MOMENTA

We have not yet specified properties of the parameter λ. The timelike geodesics of the metric are most easily specified when the particle trajectory is parameterized by the proper time τ. For $\lambda_0 = 0$ if we can solve the integral Eq. (72) for the whole range of λ_1 we find $\tau(\lambda_1)$ and can in principle invert this function to write the position as a function of τ, $x^\mu(\tau)$. The resulting velocity four-vector is so interesting that we give it special letter and refer to it as *the* four-vector velocity

$$u^\mu = \frac{dx^\mu}{d\tau}. \tag{76}$$

The most important property of the four-velocity is that it is normalised. Consider the equation (for a trajectory starting at $\tau = 0$ and ending at $\tau = \tau'$)

$$\tau' = \int_0^{\tau'} d\tau \sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}}. \tag{77}$$

We may differentiate this with respect to the final proper time τ' to obtain (by using the fundamental theorem of calculus)

$$1 = \sqrt{-g_{\mu\nu}\frac{dx^\mu}{d\tau}(\tau')\frac{dx^\nu}{d\tau}(\tau')} = -g_{\mu\nu}\frac{dx^\mu}{d\tau}(\tau')\frac{dx^\nu}{d\tau}(\tau'). \tag{78}$$

Where the final equality follows because the value on the left hand side is one. Using the convention that the metric "lowers" an index we may write this equation in the shorthand

$$u_\mu u^\mu = -1. \tag{79}$$

Now our particle will have mass. This mass can be measured by testing the response of a particle at rest to a force, we will call it m_0. It is then common do define a four-momentum in analogy with classical mechanics

$$p^\mu = m_0 u^\mu. \tag{80}$$

For null trajectories we cannot improve on $dx^\mu/d\lambda$ since there is no proper time for such a trajectory. We will defer a discussion of the appropriate momentum four-vector until we get to the special relativity section.

V. SCALAR PRODUCTS

In physics we are interested in quantities that are independent of the co-ordinate system we have used to describe our physical system. In general relativity this invariance is sometimes known as general covariance. It generalises the invariance under rotations and translations familiar from Newtonian mechanics, and under Lorentz transformations as in special relativity.

So for example we should bear in mind that the 4-vectors and tensors that appear in the theory have a co-ordinate independent existence.

Many of the most important quantities we are interested in can be written as scalar products of some kind. For two 4-vectors x^μ and y^μ the scalar product is

$$\mathbf{x}\cdot\mathbf{y} = g_{\mu\nu}x^\mu y^\nu = x_\mu y^\mu. \tag{81}$$

Scalar products of vectors do not depend on the co-ordinates we have chosen.

VI. CHANGE OF CO-ORDINATES

Speaking of changes of co-ordinates. If we change to a new set of co-ordinates $x'^{\mu}(x^{\nu})$ then we must use expressions for the old co-ordinates in terms of the new and the chain rule in all our expressions to get rid of the old co-ordinates. In particular the metric changes according to

$$g_{\mu\nu}(x') = g_{\kappa\lambda}(x(x'))\frac{\partial x^{\kappa}}{\partial x'^{\mu}}\frac{\partial x^{\lambda}}{\partial x'^{\nu}}. \tag{82}$$

Check that this formula guarantees that the scalar product of two four-velocities is invariant under co-ordinate transformations

VII. GEODESICS

In the absence of forces other than gravity our particle will follow a geodesic. The geodesics of a timelike path extremize the proper time. In special relativity and for most curved spacetimes, the proper time is in fact a local maximum. The calculus of variations allows us to follow these extremizing paths given initial values of x^{μ} and u^{μ}.

First lets introduce the notation

$$w\left(x^{\mu}, \frac{dx^{\mu}}{d\tau}\right) = -g_{\mu\nu}(x)\frac{dx^{\mu}}{d\tau}\frac{dx^{\nu}}{d\tau}. \tag{83}$$

Notice that we are interpreting the co-ordinates x^{μ} and velocities u^{μ} as *independent variables* for the purposes of determining the geodesic equations.

Geodesics are paths that extremize the proper time of the path. So for example if we have a path $x^{\mu}(\lambda)$ that connects $x_A = x(\lambda = 0)$ with $x_B = x(\lambda = 1)$ then the proper time of the path is

$$\tau_{AB} = \int_0^1 d\lambda\sqrt{-g_{\mu\nu}(x)\frac{dx^{\mu}}{d\tau}\frac{dx^{\nu}}{d\tau}} = \int_0^1 d\lambda w^{1/2}. \tag{84}$$

The geodesics satisfy the Euler-Lagrange equations and these may be written in the form

$$\frac{d}{d\tau}\frac{\partial L}{\partial(dx^{\mu}/d\tau)} = \frac{\partial L}{\partial x^{\mu}}. \tag{85}$$

where $L(x(\sigma), \dot{x}(\sigma)) = w^{1/2}$.

This gives us

$$\frac{d}{d\tau}w^{-1/2}\frac{\partial w}{\partial(dx^{\mu}/d\tau)} = w^{-1/2}\frac{\partial w}{\partial x^{\mu}}/2. \tag{86}$$

This in turn is

$$-w^{-3/2}\frac{dw}{d\tau}\left(\frac{\partial w}{\partial(dx^{\mu}/d\tau)}\right)/2 + w^{-1/2}\frac{d}{d\tau}\frac{\partial w}{\partial(dx^{\mu}/d\tau)} = w^{-1/2}\frac{\partial w}{\partial x^{\mu}}/2. \tag{87}$$

Where I have just used the product and chain rule. However at this point we should not that $dw/d\tau = 0$ since, regarded as a function of τ is a constant (1).

Therefore we get a version of the Euler-Lagrange equations

$$\frac{d}{d\tau}\frac{\partial w}{\partial(dx^{\mu}/d\tau)} = \frac{\partial w}{\partial x^{\mu}}. \tag{88}$$

This is just as if we had ignored the square root. From now on when I talk about the Euler-Lagrange or geodesic equations this is what I mean.

Notice that if the right hand side is zero then it is possible to find a quantity that is constant along geodesics by looking at the argument of the derivative on the left-hand-side.

If we take the example of flat space or special relativity then we have

$$\frac{\partial w}{\partial(dx^{\mu}/d\tau)} = 2\eta_{\mu\nu}\frac{dx^{\nu}}{d\tau} \tag{89}$$

while $\partial w/\partial x^\mu = 0$. So the geodesic equations are

$$\frac{d^2x^\mu}{d\tau^2} = 0 \tag{90}$$

and so the geodesics are straight lines as we expect.

It is also sometimes useful to have the formula

$$\frac{d^2x^\mu}{d\tau^2} = -\Gamma^\mu_{\kappa\lambda}\frac{dx^\kappa}{d\tau}\frac{dx^\lambda}{d\tau} \tag{91}$$

where the Christoffel symbols are given by

$$\Gamma_{\mu\nu\kappa} = \frac{1}{2}\left(\partial_\mu g_{\kappa\nu} + \partial_\nu g_{\kappa\mu} - \partial_\kappa g_{\mu\nu}\right) \tag{92}$$

and

$$\Gamma^\kappa_{\mu\nu} = g^{\kappa\lambda}\Gamma_{\mu\nu\lambda}. \tag{93}$$

Recall that the metric with raised indices is just the inverse of the lower index version

$$g^{\mu\kappa}g_{\kappa\nu} = \delta^\mu_\nu \tag{94}$$

VIII. PRINCIPLE OF EQUIVALENCE: TIME DILATION AND REDSHIFTS

We can use the principal to find estimate the size of the gravitational time dilation effects, by using the analogy to the same problem in a spaceship undergoing uniform acceleration.

Consider two clocks A and B. With A at the top of the rocket ship send signals to B at the bottom with signals spaced in time by $\Delta\tau_A$. The goal is to find the time interval $\Delta\tau_B$ separating the arrival of signals at B.

So if the two clocks are separated by a height h and accelerates at g we have

$$z_B(t) = gt^2/2 \quad z_A(t) = h + gt^2/2 \tag{95}$$

A emits a pulse at $t = 0$ that is recieved at t_1 and a second pulse at $\Delta\tau_A$ that is recieved at $t_1 + \Delta\tau_B$.

The distance travelled by the first pulse is

$$z_A(0) - z_B(t_1) = ct_1 \tag{96}$$

and by the second pulse

$$z_A(\Delta\tau_A) - z_B(t_1 + \Delta\tau_B) = c(t_1 + \Delta\tau_B - \Delta\tau_A). \tag{97}$$

Substituting into the first equation we find

$$h - gt_1^2/2 = ct_1 \tag{98}$$

while substituting into the second equation and keeping only the linear terms in $\Delta\tau$ we find

$$h - gt_1^2/2 - gt_1\Delta\tau_B = c(t_1 + \Delta\tau_B - \Delta\tau_A) \tag{99}$$

Subtracting these equations gives

$$-gt_1\Delta\tau_B = c(\Delta\tau_B - \Delta\tau_A) \tag{100}$$

and using $t_1 \simeq h/c$ we find

$$\Delta\tau_B = \Delta\tau_A(1 - gh/c^2) \tag{101}$$

Now in the gravitational version of the problem we find $\phi_A - \phi_B = gh$ so that if $r_{A/B}$ is the rate that signals are recieved at A/B we find

$$r_B = r_A(1 + (\phi_A - \phi_B)/c^2). \tag{102}$$

The signals in question could be the crests of an electromagnetic wave. So we should also be able to use this formula to calculate gravitational redshifts. If the signal is sent from a radius R away from a mass M ($\phi_A = -GM/R$) out to infinity this suggests the redshift is

$$\omega_\infty = (1 - \frac{GM}{Rc^2})\omega_A \tag{103}$$

This turns out to be a good approximation to the redshift that is calculated in a full description of general relativity. It says that if light is emitted in a gravitational potential well it is shifted to longer wavelengths as it climbs out of that well.

IX. LIMITS TO THE PRINCIPAL OF EQUIVALENCE

In real gravitational fields the principal of equivalence breaks down as experiments are performed over larger areas or longer times. Consider for example two balls inside a space-shuttle in a circular obit of the earth. If the radius of the orbit with respect to the center of the earth is R and on ball is a small distance s above R we know that for the nearer ball the acceleration is

$$v^2/R = GM/R^2 = g \tag{104}$$

where v is the velocity of the orbital motion. The circular orbit has period $P = 2\pi R/v$. On other hand for the further ball the acceleration is

$$GM/(R+s)^2 \simeq g - 2g\frac{s}{R} \tag{105}$$

So the relative acceleration is $a_{\rm rel} = 2gs/R$ and the resulting change in the relative displacement is $\delta s = a_{\rm rel}t^2/2$. Notice that $g/R = v^2/R^2 = (P/2\pi)^2$. As a result we find that the relative displacement after a time t is

$$\delta s/s = \left(\frac{2\pi t}{P}\right)^2 . \tag{106}$$

The message of this calculation is that the relative change in the displacement of the two balls starts to become large relative to the initial displacement when the time elapsed starts to be a significant fraction of an orbital period. This gives a timescale over which tidal effects start to be important.