General Coordinate Transformations in Minkowski Space

In order to move away from constant accelerations (as models of observers in constant gravitational fields only), we now consider the effect of arbitrary (general) coordinate transformations on the laws of special relativity and the geometry of Minkowski space. This may look like a somewhat exaggerated move at this point (should we perhaps not just look at coordinate transformations to coordinates that somehow correspond to adapted coordinates for some arbitrary accelerated observer?), but

- it is actually easier to just do this than to understand what is meant precisely by this parenthetical remark and how to implement it;
- there are many useful things that one can learn from doing this;
- and we will see later (when discussing the relation between the *Einstein Equivalence Principle* and the *Principle of General Covariance* in section 3.1), that the relation between the description of physics in an arbitrary gravitational field and the behaviour of this description under arbitrary coordinate transformations is much closer and more far-reaching than we perhaps have the right to expect at the moment.

Let us see what the equation of motion (1.51) of a free massive particle looks like when written in some other (non-inertial, accelerating) coordinate system. It is extremely useful for bookkeeping purposes and for avoiding algebraic errors to use different kinds of indices for different coordinate systems. Thus we will call the new coordinates $x^\mu(\xi^b)$ and not, say, $x^a(\xi^b)$.

First of all, proper time should not depend on which coordinates we use to describe the motion of the particle (the particle couldn't care less what coordinates we experimenters or observers use). [By the way: this is the best way to resolve the so-called 'twin-paradox': It doesn't matter which reference system you use - the accelerating twin in the rocket will always be younger than her brother when they meet again.] Thus

$$\begin{aligned} d\tau^2 &= -\eta_{ab} d\xi^a d\xi^b \\ &= -\eta_{ab} \frac{\partial \xi^a}{\partial x^\mu} \frac{\partial \xi^b}{\partial x^\nu} dx^\mu dx^\nu \quad . \end{aligned} \tag{1.78}$$

Here

$$J^a_\mu(x) = \frac{\partial \xi^a}{\partial x^\mu} \tag{1.79}$$

is the Jacobi matrix associated to the coordinate transformation $\xi^a = \xi^a(x^\mu)$, and we will make the assumption that (locally) this matrix is non-degenerate, thus has an inverse $J^\mu_a(x)$ or $J^\mu_a(\xi)$ which is the Jacobi matrix associated to the inverse coordinate transformation $x^\mu = x^\mu(\xi^a)$,

$$J^a_\mu J^\mu_b = \delta^a_b \qquad J^\mu_a J^a_\nu = \delta^\mu_\nu \ . \tag{1.80}$$

We see that in the new coordinates, proper time and distance are no longer measured by the Minkowski metric in its standard form (the constant matrix η_{ab}), but by

$$d\tau^2 = -g_{\mu\nu}(x)dx^\mu dx^\nu \ , \tag{1.81}$$

where the *metric tensor* (or *metric* for short) $g_{\mu\nu}(x)$ is

$$g_{\mu\nu}(x) = \eta_{ab}\frac{\partial\xi^a}{\partial x^\mu}\frac{\partial\xi^b}{\partial x^\nu} \ . \tag{1.82}$$

The fact that the Minkowski metric written in the coordinates x^μ in general depends on x should not come as a surprise - after all, this also happens when one writes the Euclidean metric in spherical coordinates etc.

It is easy to check, using (1.80), that the inverse metric, which we will denote by $g^{\mu\nu}$,

$$g^{\mu\nu}(x)g_{\nu\lambda}(x) = \delta^\mu_{\ \lambda} \ , \tag{1.83}$$

is given by

$$g^{\mu\nu}(x) = \eta^{ab}\frac{\partial x^\mu}{\partial\xi^a}\frac{\partial x^\nu}{\partial\xi^b} \ . \tag{1.84}$$

We will have much more to say about the metric below and, indeed, throughout this course.

Turning now to the equation of motion, the usual rules for a change of variables give

$$\frac{d}{d\tau}\xi^a = \frac{\partial\xi^a}{\partial x^\mu}\frac{dx^\mu}{d\tau} \ , \tag{1.85}$$

where $\frac{\partial\xi^a}{\partial x^\mu}$ is an invertible matrix at every point. Differentiating once more, one finds

$$\begin{aligned}\frac{d^2}{d\tau^2}\xi^a &= \frac{\partial\xi^a}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \frac{\partial^2\xi^a}{\partial x^\nu\partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} \\ &= \frac{\partial\xi^a}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} + \delta^a_{\ b}\frac{\partial^2\xi^b}{\partial x^\nu\partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} \\ &= \frac{\partial\xi^a}{\partial x^\mu}\left[\frac{d^2x^\mu}{d\tau^2} + \frac{\partial x^\mu}{\partial\xi^b}\frac{\partial^2\xi^b}{\partial x^\nu\partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \ . \end{aligned} \tag{1.86}$$

Thus, since the matrix appearing outside the square bracket is invertible, in terms of the coordinates x^μ the equation of motion, or the equation for a straight line in Minkowski space, becomes

$$\frac{d^2x^\mu}{d\tau^2}+\frac{\partial x^\mu}{\partial\xi^a}\frac{\partial^2\xi^a}{\partial x^\nu\partial x^\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}=0\ \ . \tag{1.87}$$

The second term in this equation, which we will write as

$$\frac{d^2x^\mu}{d\tau^2}+\Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}=0\ \ , \tag{1.88}$$

where

$$\Gamma^\mu_{\ \nu\lambda}=\frac{\partial x^\mu}{\partial\xi^a}\frac{\partial^2\xi^a}{\partial x^\nu\partial x^\lambda}\ \ , \tag{1.89}$$

or, more compactly,

$$\Gamma^\mu_{\ \nu\lambda}=J^\mu_a\partial_\nu J^a_\lambda=J^\mu_a\partial_\lambda J^a_\nu\equiv J^\mu_a J^a_{\nu\lambda}\ \ , \tag{1.90}$$

represents a pseudo-force or fictitious gravitational force (like a centrifugal force or the Coriolis force) that arises whenever one describes inertial motion in non-inertial coordinates. This term is absent for linear coordinate transformations $\xi^a(x^\mu)=M^a_\mu x^\mu$. In particular, this means that the equation (1.51) is invariant under Lorentz transformations, as it should be.

While (1.88) looks a bit complicated and unattractive, it is simply the general variant of a calculation that you have probably done numerous times before in various specific contexts. Moreover, there are at least two very useful things that we can extract or anticipate from this equation, namely

1. candidates for the appropriate generalisation of the Newtonian gravitational potential
2. the prototypical *general covariance* of physical equations

in *any* theory of gravity satisfying the Einstein equivalence principle. Let us now discuss these features in turn (relegating some uninspiring calculational details to the end of this subsection):

1. the Metric as a Candidate for the Gravitational Potential

   It turns out that the above (pseudo-)force term can be expressed in terms of the partial derivatives of the metric (1.82) as

$$\begin{aligned}\Gamma^\mu_{\ \nu\lambda}&=g^{\mu\rho}\Gamma_{\rho\nu\lambda}\\ \Gamma_{\rho\nu\lambda}&=\tfrac{1}{2}(g_{\rho\nu,\lambda}+g_{\rho\lambda,\nu}-g_{\nu\lambda,\rho})\end{aligned} \tag{1.91}$$

It is an elementary but nevertheless useful exercise to check this (see below - but do try this yourself as well).

This shows that the components of the metric appear to play the role of "potentials" for the gravitational pseudo-force. In particular, since in principle all components of the metric can contribute to $\Gamma_{\rho\nu\lambda}$, we learn the interesting fact that in order to achieve this a single scalar potential, as in the Newtonian theory, is completely insufficient.

If the metric indeed plays the role of the gravitational potential, as suggested by these considerations, then it will play the role of the fundamental dynamical variable of gravity. Since the metric encodes what one usually refers to as the geometry of a space(-time), namely the information required to determine distances, areas, volumes etc., this means that we are being led to the conclusion that any theory of gravity based on the equivalence principle is a theory of dynamical geometry. Wow ...

2. the General Covariance of the Equation of Motion

   The equation of motion (1.88) has one other fundamental redeeming and attractive feature which will also make it the prototype of the kind of equations that we will be looking for in general. This feature is its *covariance* under general coordinate transformations, i.e. its *general covariance*, which means that the equation takes the same form in any coordinate system. Indeed, this covariance is in some sense tautologically true since the coordinate system $\{x^\mu\}$ that we have chosen is indeed arbitrary. However, it is instructive to see how this comes about by explicitly transforming (1.88) from one coordinate system x^μ to another, say y^α.

If one does this (cf. below for a proof), one finds that the equations of motion (1.88) in the coordinates x^μ and y^α are related by

$$\frac{d^2y^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = \frac{\partial y^\alpha}{\partial x^\mu}\left[\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \tag{1.92}$$

Thus the geodesic equation transforms in the simplest possible non-trivial way under coordinate transformations $x \to y$, namely with the Jacobi matrix

$$J^\alpha_\mu = \frac{\partial y^\alpha}{\partial x^\mu} \quad . \tag{1.93}$$

We will see later that this transformation behaviour characterises/defines tensors, in this particular case a vector (or contravariant tensor of rank 1).

In particular, since this matrix is assumed to be invertible, we reach the conclusion that the left hand side of (1.92) is zero if and only if the term in square brackets on the right hand side is zero,

$$\frac{d^2y^\alpha}{d\tau^2}+\Gamma^\alpha_{\ \beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau}=0 \quad \Leftrightarrow \quad \frac{d^2x^\mu}{d\tau^2}+\Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}=0 \tag{1.94}$$

This is what is meant by the statement that the equation takes the same form in any coordinate system, and is therefore satisfied in one coordinate system if and only if it is satisfied in all coordinate systems. We see that in this case this is achieved by having the equation transform in a particularly simple way under coordinate transformations, namely as a tensor.

One might then, on the basis of the equivalence principle, also want to postulate that the motion of particles in a general gravitational field, described by a metric, is then still governed by (1.88) and (1.91). In this more general context the $\Gamma^\mu_{\ \nu\lambda}$ are referred to as the *Christoffel symbols* of the metric.

Happily, as we will see below, in section 1.7, these equations need not be postulated at all - they are simply the *geodesic equations* satisfied by paths that extremise proper time (or proper distance), and are thus the Euler-Lagrange equations for the obvious generalisation of the special relativistic action for a free particle, $S \sim \int d\tau$, to an arbitrary metric.

---

1. Proof of (1.91):

   - From

   $$g_{\mu\nu}=\eta_{ab}J^a_\mu J^b_\nu \tag{1.95}$$

   one deduces

   $$g_{\mu\nu,\lambda}=\eta_{ab}(J^a_{\mu\lambda}J^b_\nu+J^a_\mu J^b_{\nu\lambda}) \tag{1.96}$$

where

$$J^a_{\mu\lambda} = \partial_\lambda J^a_\mu = \frac{\partial^2 \xi^a}{\partial x^\mu \partial x^\lambda} = J^a_{\lambda\mu} \quad . \tag{1.97}$$

- Therefore, now adopting (1.91) as the definition of the Γ-symbols, one has

$$\begin{aligned} \Gamma_{\mu\nu\lambda} &= \tfrac{1}{2}(g_{\mu\nu,\lambda} + g_{\mu\lambda,\nu} - g_{\nu\lambda,\mu}) \\ &= \tfrac{1}{2}\eta_{ab}(J^a_{\mu\lambda}J^b_\nu + J^a_\mu J^b_{\nu\lambda} + J^a_{\mu\nu}J^b_\lambda + J^a_\mu J^b_{\lambda\nu} - J^a_{\nu\mu}J^b_\lambda - J^a_\nu J^b_{\lambda\mu}) \\ &= \eta_{ab}J^a_\mu J^b_{\nu\lambda} \quad , \end{aligned} \tag{1.98}$$

where the cancellations in passing to the last line arise from the symmetries $\eta_{ab} = \eta_{ba}$, $J^b_{\lambda\mu} = J^b_{\mu\lambda}$ etc.

- Thus, finally (and writing out everything in detail for once),

$$\begin{aligned} \Gamma^\mu{}_{\nu\lambda} = g^{\mu\rho}\Gamma_{\rho\nu\lambda} &= \eta^{cd}J^\mu_c J^\rho_d \eta_{ab} J^a_\rho J^b_{\nu\lambda} = \eta^{cd}J^\mu_c \delta^a_d \eta_{ab} J^b_{\nu\lambda} \\ &= \eta^{ca}J^\mu_c \eta_{ab} J^b_{\nu\lambda} = \delta^c_b J^\mu_c J^b_{\nu\lambda} = J^\mu_b J^b_{\nu\lambda} \quad , \end{aligned} \tag{1.99}$$

as was to be shown.

2. Proof of (1.92):

- Consider transforming the free particle equation of motion in inertial coordinates (1.51) not to the coordinate systsem x^μ, as we did before, but to another coordinate system $\{y^\alpha\}$. Following the same steps as above, one arrives at the y-version of (1.86), namely

$$\frac{d^2}{d\tau^2}\xi^a = \frac{\partial \xi^a}{\partial y^\alpha}\left[\frac{d^2 y^\alpha}{d\tau^2} + \frac{\partial y^\alpha}{\partial \xi^b}\frac{\partial^2 \xi^b}{\partial y^\beta \partial y^\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau}\right] \quad . \tag{1.100}$$

- Equating this result to (1.86) and using the chain rule for partial derivatives

$$\frac{\partial y^\alpha}{\partial x^\mu} = \frac{\partial y^\alpha}{\partial \xi^a}\frac{\partial \xi^a}{\partial x^\mu} \quad , \tag{1.101}$$

one finds

$$\frac{d^2 y^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma}\frac{dy^\beta}{d\tau}\frac{dy^\gamma}{d\tau} = \frac{\partial y^\alpha}{\partial x^\mu}\left[\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu{}_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau}\right] \tag{1.102}$$

as claimed.

1.5 Metrics I: Definition and Examples

Above we saw that the motion of free particles in Minkowski space in curvilinear coordinates is described in terms of a modified metric, $g_{\mu\nu}$, and a force term $\Gamma^\mu{}_{\nu\lambda}$ representing the 'pseudo-force' on the particle. Thus the Einstein Equivalence Principle suggests that an appropriate description of true gravitational fields is in terms of a metric tensor $g_{\mu\nu}(x)$ (and its associated Christoffel symbols) which can only locally be related to the

Minkowski metric via a suitable coordinate transformation (to locally inertial coordinates). Thus our starting point will now be a space-time equipped with some metric $g_{\mu\nu}(x)$, which (by analogy with the Euclidean and Minkowski metrics) we will assume to be symmetric and non-degenerate, i.e.

$$g_{\mu\nu}(x) = g_{\nu\mu}(x) \qquad \det(g_{\mu\nu}(x)) \neq 0 \ . \tag{1.103}$$

The metric encodes the information how to measure (spatial and temporal) distances, as well as areas, volumes etc., via the associated line element

$$g_{\mu\nu}(x) \quad \Rightarrow \quad ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \ . \tag{1.104}$$

Thus a metric determines a geometry (in the literal sense of a prescription for measuring distances etc.), but different metrics may well determine the same geometry, namely those metrics which are just related by coordinate transformations. In particular, distances should not depend on which coordinate system is used. Hence, changing coordinates from the $\{x^\mu\}$ to new coordinates $\{y^\alpha(x^\mu)\}$ and demanding that

$$g_{\mu\nu}(x)dx^\mu dx^\nu = g_{\alpha\beta}(y)dy^\alpha dy^\beta \ , \tag{1.105}$$

one finds that under a coordinate transformation a metric transforms as

$$g_{\alpha\beta}(y) = g_{\mu\nu}(x)\frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial x^\nu}{\partial y^\beta} \equiv J^\mu_\alpha J^\nu_\beta \, g_{\mu\nu} \ . \tag{1.106}$$

Objects which transform in such a nice and simple way under coordinate transformations are known as *tensors* - the metric is an example of what is known as (and we will get to know as) a covariant symmetric rank two tensor. We will study tensors in much more detail and generality later, starting in section 3.

<u>REMARKS:</u>

1. Here I have denoted the components of the metric in the new coordinates y^α simply by $g_{\alpha\beta}$. Occasionally it is more convenient to use a more elaborate notation, such as

$$x^\mu \to x'^\alpha = y^\alpha \quad \Rightarrow \quad g_{\mu\nu} \to g'_{\alpha\beta} = J^\mu_\alpha J^\nu_\beta g_{\mu\nu} \ , \tag{1.107}$$

   which allows one to distinguish notationally specific components of the metric in 2 different coordinate systems, such as g'_{11} (the (11)-component of the metric in the y-coordinates) from g_{11} (the (11)-component of the metric in the x-coordinates). As mentioned before, indices and other decorations are primarily bookkeeping devices; therefore I will usually not be overly-pedantic about these things in the following and will use whatever notation is more convenient in the case at hand.

2. As a consequence of the non-degeneracy condition, pointwise $g_{\mu\nu}(x)$ possesses an inverse, whose components we will denote by $g^{\mu\nu}(x)$, i.e.

$$g^{\mu\nu}(x)g_{\nu\lambda}(x) = \delta^{\mu}_{\lambda} \quad , \quad g_{\mu\nu}(x)g^{\nu\lambda}(x) = \delta^{\lambda}_{\mu} \ . \tag{1.108}$$

Clearly, the inverse metric then transforms inversely, i.e. with the inverse Jacobi matrices J^{α}_{μ}, and this is now nicely compatible with the convention to denote the inverse metric by upper indices,

$$g^{\alpha\beta} = J^{\alpha}_{\mu} J^{\beta}_{\nu} g^{\mu\nu} \ . \tag{1.109}$$

This is also the rationale for writing the invese metric with "upper" indices: the positioning of indices is used to indicate how an object transforms under coordinate transformations (and we will formalise this in the discussion of section 3 on tensor algebra).

3. A space-time equipped with a metric tensor $g_{\mu\nu}(x)$ is called a metric space-time or (pseudo-)Riemannian space-time. Here "Riemannian" usually refers to a space equipped with a positive-definite metric (all eigenvalues positive), while pseudo-Riemannian (or Lorentzian) refers to a space-time with a metric with one negative and 3 (or 27, or whatever) positive eigenvalues.

4. One point to note about the tensorial transformation behaviour is that pointwise it is a similarity transformation in the sense of linear algebra, in matrix notation

$$g \mapsto J^t g J \ . \tag{1.110}$$

In particular, therefore, if in one coordinate system the space-time metric tensor has one negative and three positive eigenvalues (as in a locally inertial coordinate system), then the same will be true in any other coordinate system (even though the eigenvalues themselves will in general be different) - this statement should be familiar from linear algebra (e.g. as Sylvester's law of inertia, but it also goes under various other names).

Here are some examples of Riemannian metrics that you may already be familiar with.

<u>EXAMPLES:</u>

1. The Euclidean metrics or line-elements on $\mathbb{R}^2$ or $\mathbb{R}^3$, but written in polar or spherical coordinates,

$$\begin{aligned} ds^2(\mathbb{R}^2) &= dx^2 + dy^2 = dr^2 + r^2 d\phi^2 \\ ds^2(\mathbb{R}^3) &= dx^2 + dy^2 + dz^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \end{aligned} \tag{1.111}$$

E.g. for the latter case one has

$$(x, y, z) = (r \sin\theta \cos\phi, r \sin\theta \sin\phi, r \cos\theta) \ , \tag{1.112}$$

and plugging this into the Euclidean line-element $dx^2 + dy^2 + dz^2$, one finds the above result.

Denoting the Cartesian coordinates by x^α and the spherical coordinates by y^α, with $(y^1 = r, y^2 = \theta, y^3 = \phi)$, the non-vanishing components of the metric in the two coordinate systems are thus (using the prime notation (1.107))

$$g_{11} = g_{22} = g_{33} = 1 \quad , \quad g'_{11} = 1 \ , \ g'_{22} = r^2 \ , \ g'_{33} = r^2 \sin^2\theta \ . \tag{1.113}$$

Alternatively, it is often more informative (and very common) to use the coordinates themselves, rather than indices, as the labels of the components of the metric tensor. In this case one can dispense with the prime notation and simply write the components of the metric in spherical coordinates as

$$g_{rr} = 1 \ , \ g_{\theta\theta} = r^2 \ , \ g_{\phi\phi} = r^2 \sin^2\theta \ . \tag{1.114}$$

2. Restricting the first example above to constant radius $r = R$, this gives us the line-element on the circle S^1_R of radius R,

$$ds^2(S^1_R) = R^2 d\phi^2 \ . \tag{1.115}$$

Restricting the second to the 2-sphere S^2_R of radius R,

$$x^2 + y^2 + z^2 = r^2 = R^2 \quad \text{or} \quad r = R \ , \tag{1.116}$$

one finds the line-element

$$ds^2(S^2_R) = R^2(d\theta^2 + \sin^2\theta d\phi^2) \equiv R^2 d\Omega^2 \ . \tag{1.117}$$

Here

$$d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2 \tag{1.118}$$

is usually called the solid angle, and we can now interpret it as the line element on the unit 2-sphere. We will use the notation / abbrevation $d\Omega^2$ for this line element throughout the notes.

This example provides a nice illustration of the fact that by drawing the coordinate grid / infinitesimal parallelograms determined by the metric tensor, one can get a feeling for the geometry and can in particular convince onseself that in general a metric space or space-time need not or cannot be flat, i.e. is not the flat Euclidean space of Euclidean geometry.

Indeed, the coordinate grid of the metric $d\theta^2 + \sin^2\theta d\phi^2$ cannot be drawn in flat space because the infinitesimal parallelograms described by ds^2 degenerate to triangles not just at $\theta = 0$ (as would also be the case for the flat metric $ds^2 = dr^2 + r^2 d\phi^2$ in polar coordinates at $r = 0$), but also at $\theta = \pi$. This coordinate grid can, on the other hand, of course be drawn on the 2-sphere.

3. This line-element on the unit 2-sphere generalises to the line-element on a unit 3-sphere,

$$ds^2(S^3) = d\psi^2 + \sin^2\psi(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{1.119}$$

This can be obtained by simply generalising the construction of spherical coordinates from $\mathbb{R}^3$ to $\mathbb{R}^4$, and (if required) this can be continued iteratively to yet higher-dimensional spheres.

Alternatively, by thinking of the 3-sphere as the locus

$$x^2 + y^2 + z^2 + w^2 = 1 \tag{1.120}$$

in $\mathbb{R}^4$, and "solving" this equation by first setting

$$x^2 + y^2 = \sin^2\alpha \quad , \quad z^2 + w^2 = \cos^2\alpha \quad , \tag{1.121}$$

and then refining this to

$$x = \sin\alpha\cos\beta \quad , \quad y = \sin\alpha\sin\beta \quad , \quad z = \cos\alpha\cos\gamma \quad , \quad w = \cos\alpha\sin\gamma \quad , \tag{1.122}$$

one finds that the standard Euclidean line-element

$$ds^2 = dx^2 + dy^2 + dz^2 + dw^2 \tag{1.123}$$

induces the line-element

$$ds^2(S^3) = d\alpha^2 + \sin^2\alpha d\beta^2 + \cos^2\alpha d\gamma^2 \tag{1.124}$$

on the sphere. This is the same metric on S^3 as above (1.119), namely the one induced from the Euclidean metric on $\mathbb{R}^4$, but written in different coordinates. In particular, both are invariant under 4-dimensional rotations, i.e. under $SO(4)$-transformations.

However, we can obtain genuinely different metrics on the 3-sphere e.g. by starting with different metrics on $\mathbb{R}^4$. One of the simplest possiblilities is to replace (1.123) by

$$d\tilde{s}^2 = a^2(dx^2 + dy^2) + b^2(dz^2 + dw^2) \quad , \tag{1.125}$$

with a, b real non-zero parameters. Then the induced metric on the 3-sphere $x^2 + y^2 + z^2 + w^2 = 1$ is

$$d\tilde{s}^2(S^3) = (a^2\cos^2\alpha + b^2\sin^2\alpha)d\alpha^2 + a^2\sin^2\alpha d\beta^2 + b^2\cos^2\alpha d\gamma^2 \quad . \tag{1.126}$$

For $a^2 \neq b^2$, this metric is not invariant under full 4-dimensional rotations, but only under rotations in the (x, y) and (z, w) planes, i.e. under $SO(2) \times SO(2)$ transformations. Thus this equips the 3-sphere with a genuinely different geometry (and is an example of what is sometimes referred to as a "squashed 3-sphere geometry").

4. If instead of the unit 2-sphere one considers the "unit" hyperboloid H^2, defined by
$$x^2 + y^2 + z^2 = +1 \quad \longrightarrow \quad x^2 + y^2 - z^2 = -1 \ , \tag{1.127}$$
then this is naturally thought of as being embedded not in $\mathbb{R}^3$ but in $\mathbb{R}^{1,2}$, i.e. into the 3-dimensional vector space with line-element
$$ds^2 = dx^2 + dy^2 - dz^2 \ . \tag{1.128}$$
The hyperbolic analogues (r, σ, ϕ) of the spherical coordinates, defined by
$$(x, y, z) = (r \sinh\sigma \cos\phi, r \sinh\sigma \sin\phi, r \cosh\sigma) \ , \tag{1.129}$$
are naturally adapted to this situation, because
$$x^2 + y^2 - z^2 = -r^2 \tag{1.130}$$
so that the unit hyperboloid is evidently just the surface $r = 1$. In these coordinates, the metric (1.128) takes the form
$$ds^2 = -dr^2 + r^2(d\sigma^2 + \sinh^2\sigma \ d\phi^2) \ , \tag{1.131}$$
and therefore the induced metric on the unit hyperboloid $r = 1$ is
$$ds^2(H^2) = d\sigma^2 + \sinh^2\sigma \ d\phi^2 \ . \tag{1.132}$$

Metrics II: Lorentzian (Pseudo-Riemannian) Metrics

We now turn to Lorentzian (pseudo-Riemannian) metrics and geometries. These will of course occupy and accompany us throughout these notes, so this section is meant to just provide a first brief encounter with these objects.

For a metric with Lorentzian signature, and with coordinates $x^\alpha = (x^0 = t, x^k)$, say, the metric has components $g_{00}, g_{0k} = g_{k0}$ and $g_{ik} = g_{ki}$, and the corresponding line element has the form
$$ds^2 = g_{00}dt^2 + 2g_{0k}dt\, dx^k + g_{ik}dx^i dx^k \tag{1.133}$$

Her are some simple

<u>Examples:</u>

1. Of course any of the Riemannian metrics of the previous section be promoted to space-time metrics by simply adding a $(-dt^2)$ (i.e. by taking the direct product with the time-axis). Thus the Minkowski metric in spatial spherical coordinates has the form
$$ds^2(\mathbb{R}^{1,3}) = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{1.134}$$

2. A generalisation of this is provided by the so-called *ultrastatic* metrics, i.e. metrics that are just a product of the standard metric $-dt^2$ along the time-direction and a spatial metric $\tilde{g}_{ij}(x)$
$$ds^2 = -dt^2 + \tilde{g}_{ij}(x)dx^i dx^j \tag{1.135}$$
(i.e. the components depend only on the spatial coordinates x^i, not on t).

3. Somewhat more generally, the spatial comopnents of the metric can depend non-trivially on time. For example, a space-time metric describing a spatially spherical universe with a time-dependent radius (expansion of the universe!) might be described by the line element
$$ds^2 = -dt^2 + a(t)^2 \left(d\psi^2 + \sin^2\psi(d\theta^2 + \sin^2\theta d\phi^2)\right) \quad , \tag{1.136}$$
and more generally one can consider the corresponding generalisation of (1.135), namely metrics of the form
$$ds^2 = -dt^2 + a(t)^2\tilde{g}_{ij}(x)dx^i dx^j \quad . \tag{1.137}$$
This describes a space-time with spatial metric $\tilde{g}_{ij}(x)dx^i dx^j$ and a time-dependent radius $a(t)$; in particular, such a space-time metric can describe an expanding universe in cosmology. We will discuss such metrics in detail later on in the context of cosmology, sections 32-37.

4. Also, the (time-time)-component of the metric can of course in general depend non-trivially on the spatial coordinates. A prominent example is the *Schwarzschild Metric*
$$ds^2 = -(1-2m/r)dt^2 + (1-2m/r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad . \tag{1.138}$$
It is of fundamental importance for General Relativity, and perhaps the most important exact solution of the Einstein field equations for the gravitational field, as it describes the gravitational field outside a spherical star (as well as black holes, as it turns out ...). We will discuss this metric in great detail in sections 23-26.

The characteristic feature of metrics with Lorentzian signature is of course the presence of timelike and null (lightlike) directions, and thus in a pseudo-Riemannian space-time one has the same distinction between spacelike, timelike and lightlike separations as in Minkowski space(-time). Infinitesimal

- spacelike distances correspond to $ds^2 > 0$,
- timelike distances to $d\tau^2 = -ds^2 > 0$,
- and null or lightlike distances to $ds^2 = d\tau^2 = 0$.

Likewise, a vector $V^\mu(x)$ at a point x is called

- spacelike if $g_{\mu\nu}(x)V^\mu(x)V^\nu(x) > 0$,
- timelike if $g_{\mu\nu}(x)V^\mu(x)V^\nu(x) < 0$,
- and null or lightlike if $g_{\mu\nu}(x)V^\mu(x)V^\nu(x) = 0$,

and a curve $x^\mu(\lambda)$ is called spacelike if its tangent vector is everywhere spacelike etc.

Using the definition of a vector in general relativity (to be introduced in section 3), namely an object that transforms in the obvious way, with the Jacobi matrix, under coordinate transformations, one sees that $g_{\mu\nu}(x)V^\mu(x)V^\nu(x)$ is a scalar, i.e. invariant under coordinate transformations, and hence the statement that a vector is, say, spacelike is a coordinate-independent statement, as it should be.

When the metric (1.133) is (time-space) block-diagonal, i.e. when the mixed components $g_{0k} = 0$ (as in all of the above examples), then the timelike and spacelike directions are easy to distinguish by inspection. Typically then the "spatial" metric g_{ik} is positive definite, and thus necessarily $g_{00} < 0$.

When some of the g_{0k} are non-zero, on the other hand, one has a more intricate mixing of time- and space-directions. This can also be seen from the components of the inverse metric. Indeed, from (1.108), one finds

$$g_{0\nu}g^{\nu k} = g_{00}g^{0k} + g_{0i}g^{ik} = \delta_0^k = 0 \quad , \tag{1.139}$$

and thus (for $g_{00} \neq 0$)

$$g^{0k} = -\frac{1}{g_{00}} g_{0i}g^{ik} \quad . \tag{1.140}$$

Likewise from (1.108) one deduces

$$g_{i\nu}g^{\nu k} = g_{i0}g^{0k} + g_{ij}g^{jk} = \delta_i^k \quad . \tag{1.141}$$

In particular, this shows that in general (i.e. unless the off-diagonal components g_{0k} are all zero), the spatial components g^{ik} of the inverse metric are not the inverse of the spatial components g_{ij} of the metric. Rather, using (1.140) one has

$$\left(g_{ij} - \frac{1}{g_{00}} g_{i0}g_{j0}\right) g^{jk} = \delta_i^k \quad . \tag{1.142}$$

This ends our first brief encounter with metrics and geometries. At this point the question naturally arises how one can tell whether a given (perhaps complicated looking) metric is just the "flat" (Euclidean or Minkowski) metric written in other coordinates or whether it describes a genuinely curved space-time. We will see later that there is an object, the *Riemann curvature tensor*, constructed from the metric and its 1st and 2nd derivatives, which has the property that all of its components vanish if and only if the metric is a coordinate transform of the flat space Minkowski metric. Thus, given a metric, by calculating its curvature tensor one can decide if the metric is just the flat metric in disguise or not. The curvature tensor will be introduced in section 7, and the above statement will be established in section 10.2.

Geodesic Equation from the Extremisation of Proper Time

We have seen that the equation for a straight line in Minkowski space, written in arbitrary coordinates, is

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\ \nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0 \ , \tag{1.143}$$

where the pseudo-force term $\Gamma^\mu_{\ \nu\lambda}$ is given by (1.89). We have also seen in (1.91) (provided you checked this) that $\Gamma^\mu_{\ \nu\lambda}$ can be expressed in terms of the metric (1.82) as

$$\Gamma^\mu_{\ \nu\lambda} = \tfrac{1}{2}g^{\mu\rho}(g_{\rho\nu,\lambda} + g_{\rho\lambda,\nu} - g_{\nu\lambda,\rho}) \ . \tag{1.144}$$

This gravitational force term is fictitious since it can globally be transformed away by going to the global inertial coordinates ξ^a. The equivalence principle suggests, however, that in general the equation for the worldline of a massive particle, i.e. a path that extremises proper time, in a true gravitational field is also of the above form.

We will now confirm this by deriving the equations for a timelike path that extremises proper time from a variational principle. These paths will be referred to as (timelike) *geodesics*. We will briefly return below to the (delicate) issue to which extent these can be regarded as world lines of actual massive particles.

Recall first of all from special relativity that the Lorentz-covariant description of the dynamics of a massive particle is based on describing the timelike worldline of the particle in the parametric form

$$\xi^a = \xi^a(\tau) \tag{1.145}$$

where τ is the proper time along the worldline,

$$d\tau^2 = -\eta_{ab}d\xi^a d\xi^b \ . \tag{1.146}$$

In particular, the 4-velocity

$$u^a = \frac{d\xi^a(\tau)}{d\tau} \tag{1.147}$$

is normalised as

$$\eta_{ab}u^a u^b = -1 \ . \tag{1.148}$$

The Lorentz-invariant action for a free massive particle with mass m is

$$S_0 = -m \int d\tau \quad . \tag{1.149}$$

We can adopt the same set-up and action in the present setting. Thus we parametrise the worldlines by

$$x^\mu = x^\mu(\tau) \quad , \tag{1.150}$$

with τ the proper time

$$d\tau^2 = -g_{\mu\nu}(x)dx^\mu dx^\nu \quad , \tag{1.151}$$

invariant under general coordinate transformations (provided that one transforms the metric appropriately). The corresponding 4-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} \tag{1.152}$$

is again normalised as

$$g_{\mu\nu}u^\mu u^\nu = -1 \quad , \tag{1.153}$$

and we are led to consider the coordinate-invariant Lagrangian

$$S_0[x] = -m \int d\tau = -m \int \sqrt{-g_{\mu\nu}(x)dx^\mu dx^\nu} \quad . \tag{1.154}$$

Of course m drops out of the variational equations (as it should by the equivalence principle) and we will therefore ignore m in the following.

In order to perform the variation, it is useful to introduce an arbitrary auxiliary parameter λ in the initial stages of the calculation via

$$d\tau = (-g_{\mu\nu}\tfrac{dx^\mu}{d\lambda}\tfrac{dx^\nu}{d\lambda})^{1/2}d\lambda \quad , \tag{1.155}$$

and to write

$$\int d\tau = \int (d\tau/d\lambda)d\lambda = \int (-g_{\mu\nu}\tfrac{dx^\mu}{d\lambda}\tfrac{dx^\nu}{d\lambda})^{1/2}d\lambda \quad . \tag{1.156}$$

We are varying the paths

$$x^\mu(\tau) \to x^\mu(\tau) + \delta x^\mu(\tau) \tag{1.157}$$

keeping the end-points fixed, and will denote the τ-derivatives by $\dot{x}^\mu(\tau)$. Under this variation, the metric $g_{\mu\nu}(x)$ varies as

$$\delta g_{\mu\nu} = g_{\mu\nu,\lambda}\,\delta x^\lambda \quad . \tag{1.158}$$

By the standard variational procedure one then finds, first of all,

$$\delta \int d\tau = \tfrac{1}{2}\int (-g_{\mu\nu}\tfrac{dx^\mu}{d\lambda}\tfrac{dx^\nu}{d\lambda})^{-1/2}d\lambda \left[-\delta g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} - 2g_{\mu\nu}\frac{d\delta x^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right] \quad . \tag{1.159}$$

Already at this stage we can revert from λ to τ, and the expression simplifies to

$$\delta \int d\tau = \tfrac{1}{2}\int d\tau \left[-(\delta g_{\mu\nu})\dot{x}^\mu\dot{x}^\nu - 2g_{\mu\nu}\frac{d\delta x^\mu}{d\tau}\dot{x}^\nu\right] \quad . \tag{1.160}$$

Integration by parts of the 2nd term (in order to eliminate the derivative of the variation) and use of (1.158) then leads to

$$\begin{aligned}\delta\int d\tau &= \frac{1}{2}\int d\tau\left[-g_{\mu\nu},_\lambda\dot{x}^\mu\dot{x}^\nu\delta x^\lambda + 2g_{\mu\nu}\ddot{x}^\nu\delta x^\mu + 2g_{\mu\nu},_\lambda\dot{x}^\lambda\dot{x}^\nu\delta x^\mu\right]\\ &= \int d\tau\left[g_{\mu\nu}\ddot{x}^\nu + \tfrac{1}{2}(g_{\mu\nu},_\lambda + g_{\mu\lambda},_\nu - g_{\nu\lambda},_\mu)\dot{x}^\nu\dot{x}^\lambda\right]\delta x^\mu\end{aligned} \tag{1.161}$$

after a suitable relabelling of the indices.

If we now adopt the definition (1.144) for an arbitrary metric,

$$\Gamma^\mu{}_{\nu\lambda} = \tfrac{1}{2}g^{\mu\rho}(g_{\rho\nu},_\lambda + g_{\rho\lambda},_\nu - g_{\nu\lambda},_\rho)\ , \tag{1.162}$$

we can write the result as

$$\delta\int d\tau = \int d\tau g_{\mu\nu}(\ddot{x}^\nu + \Gamma^\nu{}_{\rho\lambda}\dot{x}^\rho\dot{x}^\lambda)\delta x^\mu\ . \tag{1.163}$$

Thus we see that indeed the equations for a timelike geodesic in an arbitrary gravitational field are

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu{}_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0\ . \tag{1.164}$$

REMARKS:

1. By definition, massive test particles are those particles that satisfy the above geodesic equation, i.e. that follow timelike geodesics in space-time. However, it needs to be borne in mind that this notion of a test particle is a fiction, in particular as it neglects the backreaction, i.e. the change in the background gravitational field due to the mass of the particle. Moreover, real particles either have a finite extent (in which case this finite size should play a role in their equations of motion) or are considered to be point-like. However, the notion of a point-like particle is extremely dangerous and delicate in general relativity: as we will see later, if a given total mass is concentrated in a sufficiently small region of space-time (and "point-like" certainly qualifies as "sufficiently small"), then one will end up with a black hole rather than with the description of a particle. The correct description of point particles in general relativity is a complicated issue and an active area of research.[5]

2. One can also consider spacelike paths that extremise (minimise) proper distance, by using the action

$$S_0 \sim \int ds \tag{1.165}$$

[5] See e.g. E. Poisson, A. Pound, I. Vega, *The Motion of Point Particles in Curved Spacetime*, for a detailed discussion and many references (but you will need to acquire a solid understanding of tensor analysis first).

where

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{1.166}$$

is the proper distance (or arc-length in the traditional terminology of the differnetial geometry of curves).

One should also consider massless particles, whose worldlines will be null (or lightlike) paths. However, in that case one can evidently not use proper time or proper distance, since these are by definition zero along a null path, $ds^2 = 0$. We will come back to this special case, and a unified description of the massive and massless case, below (section 2.1). In all cases, we will refer to the resulting paths as *geodesics*. If required, we add the qualifier "timelike", "spacelike" or "null", and this is meaningful and unambiguous since, as we will see below, a geodesic that is initially timelike will always remain timelike etc.

We will have much more to say about geodesics and variational principles in section 2.

Christoffel Symbols and Coordinate Transformations

The Christoffel symbols play the role of the gravitational force term, and thus in this sense the components of the metric play the role of the gravitational potential. These Christoffel symbols play an important role not just in the geodesic equation but, as we will see later on, more generally in the definition of a covariant derivative operator and the construction of the curvature tensor, and thus ultimately also in the generally covariant description of the dynamics of the gravitational field itself.

Two elementary important properties of the Christoffel symbols are that they are symmetric in the second and third indices,

$$\Gamma_{\mu\nu\lambda} = \Gamma_{\mu\lambda\nu} \quad , \quad \Gamma^\mu_{\ \nu\lambda} = \Gamma^\mu_{\ \lambda\nu} \tag{1.167}$$

(this follows simply from the definition), and that symmetrising $\Gamma_{\mu\nu\lambda}$ over the first pair of indices one finds

$$\Gamma_{\mu\nu\lambda} + \Gamma_{\nu\mu\lambda} = g_{\mu\nu,\lambda} \tag{1.168}$$

(and this follows from noting that 4 of the 6 partial derivative terms of the metric cancel in this linear combination while 2 add up)

Knowing how the metric transforms under coordinate transformations, we can now also determine how the Christoffel symbols (1.144) and the geodesic equation transform. A straightforward but not particularly inspiring calculation (which you should nevertheless do) shows that under $x^\mu \to y^\alpha$ the Christoffel symbols are related by

$$\Gamma^\alpha_{\ \beta\gamma} = \Gamma^\mu_{\ \nu\lambda} \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial y^\beta} \frac{\partial x^\lambda}{\partial y^\gamma} + \frac{\partial y^\alpha}{\partial x^\mu} \frac{\partial^2 x^\mu}{\partial y^\beta \partial y^\gamma} \quad , \tag{1.169}$$

or

$$\Gamma^{\alpha}_{\ \beta\gamma} = J^{\alpha}_{\mu} J^{\nu}_{\beta} J^{\lambda}_{\gamma} \Gamma^{\mu}_{\ \nu\lambda} + J^{\alpha}_{\mu} \partial_{\beta} J^{\mu}_{\gamma} \quad . \tag{1.170}$$

Thus, $\Gamma^{\mu}_{\ \nu\lambda}$ transforms inhomogenously under coordinate transformations. If only the first term on the right hand side were present, then $\Gamma^{\mu}_{\ \nu\lambda}$ would be a tensor. However, the second term is there precisely to compensate for the fact that $\ddot{x}^{\mu}$ is also not a tensor - the combined geodesic equation turns out to transform in a nice way under coordinate transformations.

Namely, after another not terribly inspiring calculation (which you should nevertheless also do at least once in your life) , one finds

$$\frac{d^2 y^{\alpha}}{d\tau^2} + \Gamma^{\alpha}_{\ \beta\gamma} \frac{dy^{\beta}}{d\tau} \frac{dy^{\gamma}}{d\tau} = \frac{\partial y^{\alpha}}{\partial x^{\mu}} \left[\frac{d^2 x^{\mu}}{d\tau^2} + \Gamma^{\mu}_{\ \nu\lambda} \frac{dx^{\nu}}{d\tau} \frac{dx^{\lambda}}{d\tau} \right] \quad . \tag{1.171}$$

This is analogous to the result (1.92) that we had obtained before in Minkowski space, and the same remarks about covariance and tensors etc. apply. An explicit proof of (1.170) and (1.171) is given at the end of this subsection. A more general result along these lines will be established in section 4.1 below, when we introduce the covariant derivative of a vector field.

<u>Remarks:</u>

1. That the geodesic equation transforms in this simple way (namely as a vector) should not come as a surprise. We obtained this equation as a variational equation. The Lagrangian itself is a scalar (invariant under coordinate transformations), and the variation δx^{μ} is (i.e. transforms like) a vector,

   $$\delta y^{\alpha} = \frac{\partial y^{\alpha}}{\partial x^{\mu}} \delta x^{\mu} = J^{\alpha}_{\mu} \delta x^{\mu} \quad . \tag{1.172}$$

   Putting these pieces together, one finds the desired result.

2. General covariance, i.e. form-invariance under general coordinate transformations, as exhibited e.g. by the geodesic equation, is of course a desirable feature regardless of whether or not one is attempting to describe gravity. After all, the particle could not care less which coordinates we use to describe its motion, and therefore we should also formulate the equations of motion for a particle in a way that does not single out some preferred coordinate system or class of coordinate systems. This is precisely what is achieved by general covariance.

   However, here general covariance seems to have arisen somewhat coincidentally and spontaneously, and the relation between general covariance and gravity, or general covariance and the equivalence principle, may still appear to be somewhat mysterious at this point. The precise relation between the two concepts will be explained in section 3.1.

3. There is of course a very good physical reason for why the force term in the geodesic equation (quadratic in the 4-velocities) is not tensorial. This simply reflects the equivalence principle that locally, at a point (or in a sufficiently small neighbourhood of a point) you can eliminate the gravitational force by going to a freely falling (inertial) coordinate system. This would not be possible if the gravitational force term in the equation of motion for a particle were tensorial.

---

1. Proof of (1.170)

   For partial derivatives one has the chain rule $\partial_\gamma = J^\lambda_\gamma \partial_\lambda$ ("∂_λ is a covector"). Therefore for the partial derivatives of the metric one has

$$g_{\alpha\beta,\gamma} = (J^\mu_\alpha J^\nu_\beta g_{\mu\nu})_{,\gamma} = g_{\mu\nu,\lambda} J^\mu_\alpha J^\nu_\beta J^\lambda_\gamma + (J^\mu_{\alpha\gamma} J^\nu_\beta + J^\mu_\alpha J^\nu_{\beta\gamma}) g_{\mu\nu} \quad . \tag{1.173}$$

   Adding up the 3 terms comprising the Christoffel symbol $\Gamma_{\alpha\beta\gamma}$, one obtains

$$\begin{aligned} 2\Gamma_{\alpha\beta\gamma} &= g_{\alpha\beta,\gamma} + g_{\alpha\gamma,\beta} - g_{\beta\gamma,\alpha} \\ &= 2J^\mu_\alpha J^\nu_\beta J^\lambda_\gamma \Gamma_{\mu\nu\lambda} \\ &\quad + (J^\mu_{\alpha\gamma} J^\nu_\beta + J^\mu_\alpha J^\nu_{\beta\gamma} + J^\mu_{\alpha\beta} J^\nu_\gamma + J^\mu_\alpha J^\nu_{\gamma\beta} - J^\mu_{\beta\alpha} J^\nu_\gamma - J^\mu_\beta J^\nu_{\gamma\alpha}) g_{\mu\nu} \quad . \end{aligned} \tag{1.174}$$

   In the last line, the 3rd term cancels against the 5th (because $J^\mu_{\alpha\beta}$ is symmetric), the 1st term cancels against the 6th (because $J^\mu_{\alpha\gamma}$ and $g_{\mu\nu}$ are symmetric), while the 2nd and 4th term add up, so that one finds

$$\Gamma_{\alpha\beta\gamma} = J^\mu_\alpha J^\nu_\beta J^\lambda_\gamma \Gamma_{\mu\nu\lambda} + J^\mu_\alpha J^\nu_{\beta\gamma} g_{\mu\nu} \quad . \tag{1.175}$$

   Now the hard work has been done. Raising the 1st index of the Christoffel symbol, using the inverse metric

$$g^{\alpha\delta} = g^{\sigma\rho} J^\alpha_\sigma J^\delta_\rho \quad , \tag{1.176}$$

   it is now simple to see that one obtains the claimed result (4),

$$\Gamma^\alpha_{\ \beta\gamma} = g^{\alpha\delta}\Gamma_{\delta\beta\gamma} = J^\alpha_\mu J^\nu_\beta J^\lambda_\gamma \Gamma^\mu_{\ \nu\lambda} + J^\alpha_\mu J^\mu_{\beta\gamma} \quad . \tag{1.177}$$

   For example, for the 2nd term one has (just using properties of inverse Jacobi matrices and metrics)

$$\begin{aligned} g^{\alpha\delta} J^\mu_\delta J^\nu_{\beta\gamma} g_{\mu\nu} &= g^{\sigma\rho} J^\alpha_\sigma J^\delta_\rho J^\mu_\delta J^\nu_{\beta\gamma} g_{\mu\nu} = g^{\sigma\rho} J^\alpha_\sigma \delta^\mu_\rho J^\nu_{\beta\gamma} g_{\mu\nu} \\ &= g^{\sigma\mu} J^\alpha_\sigma J^\nu_{\beta\gamma} g_{\mu\nu} = \delta^\sigma_\nu J^\alpha_\sigma J^\nu_{\beta\gamma} = J^\alpha_\nu J^\nu_{\beta\gamma} \end{aligned} \tag{1.178}$$

2. Proof of (1.171)

   The 4-velocities transform as vectors (the chain rule again), $\dot{y}^\alpha = J^\alpha_\mu \dot{x}^\mu$. Therefore for the acceleration one has

$$\ddot{y}^\alpha = J^\alpha_\mu \ddot{x}^\mu + J^\alpha_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \quad . \tag{1.179}$$

Therefore

$$\begin{aligned}\ddot{y}^\alpha + \Gamma^\alpha_{\beta\gamma}\dot{y}^\beta\dot{y}^\gamma &= J^\alpha_\mu(\ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}J^\nu_\beta J^\lambda_\gamma\dot{y}^\beta\dot{y}^\gamma) + J^\alpha_\mu J^\mu_{\beta\gamma}\dot{y}^\beta\dot{y}^\gamma + J^\alpha_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \\ &= J^\alpha_\mu(\ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda) + (J^\alpha_\mu J^\mu_{\beta\gamma} + J^\alpha_{\mu\nu}J^\mu_\beta J^\nu_\gamma)\dot{y}^\beta\dot{y}^\gamma\end{aligned} \qquad (1.180)$$

The 1st term will give us the desired result, and cooperatively the 2nd term is identically zero because (use $\partial_\gamma = J^\nu_\gamma\partial_\nu$ again)

$$0 = (\delta^\alpha_\beta)_{,\gamma} = (J^\alpha_\mu J^\mu_\beta)_{,\gamma} = J^\alpha_{\mu\nu}J^\nu_\gamma J^\mu_\beta + J^\alpha_\mu J^\mu_{\beta\gamma} \;\; . \qquad (1.181)$$

Apology and Outlook

You may feel that, after a promising start in sections 1.1 and 1.3, the things that we have done subsequently, in particular in sections 1.4 and 1.8, look terribly messy. I agree, indeed they are! However, I can assure you that things will improve dramatically rather quickly and that this section 1 is by far the messiest part of the entire lecture notes.

Indeed, the main purpose and benefit of developing tensor calculus in the next couple of sections is to develop a formalism in the framework of which (among other things)

- one can avoid having to deal explicitly with objects that transform in complicated ways under coordinate transformations
- the transformation behaviour of any object is manifest (and does not have to be checked)
- it is straightforward to write down equations that are *generally covariant*, i.e. independent of the coordinate system in the sense that they are satisfied in all coordinate systems if and only if they are satisfied in one.

This tensor calculus formalism is simple, elegant and efficient and will then allow us to make rapid progress towards describing the dynamics in a (and subsequently of the) gravitational field in a way compatible with the Einstein equivalence principle.