Astronomy 3/4 – General Relativity I – Part 3
Manifolds, vectors and differentiation

Norman Gray

Martinmas Term, 2004

[Public version 1.0]

Aims You should

1. understand the relationship between directional derivatives on a manifold and the vectors in the tangent plane;

2. understand the relationship between covariant differentiation in flat and curved spaces;
3. understand the significance of the geodesic equation;

Objectives You should be able to demonstrate that you can

1. obtain the components of the basis transformation matrix Λ, given the algebraic relationship between the corresponding coordinates;
2. calculate expressions for the connection coefficients $\Gamma^i{}_{jk}$, given the metric (though you do not need to memorise Eq. (3.26)), and use these to calculate the components of covariant derivatives;
3. quote the expressions Eq. (3.15) and Eq. (3.20) for the covariant derivatives of vectors and one-forms;
4. explain why parallel transport and the connection are important in the definition of differentiation on a curved space (that is, give a summary account of the argument in Sect. 3.2);
5. use the geodesic equation (in its differential equation form, Eq. (3.34)) to obtain equations for geodesics in coordinate form, in simple cases, and for given metrics or connection coefficients;
6. calculate the Riemann and Ricci tensors corresponding to a given metric or set of connection coefficients, in simple cases (note that 'in simple cases' is important: this task is generally algebraically unwieldy unless the problem is constructed to have a good deal of symmetry, and unless you can recognise and exploit this – it is this latter ability that is the real objective here);
7. apply the geodesic deviation equation, Eq. (3.51) in simple cases.

1 The tangent vector

In the previous part, we carefully worked out the various things we can do with a set of vectors, one-forms and tensors, once we have identified those objects. Identifying those objects on a curved surface is precisely what we are about to do now. We discover that we have to take a rather roundabout route to them.

1.1 Manifolds and functions

The arena on which everything happens is the *manifold*, which is a very primitive concept. A manifold is a set of points, with the only extra structure being enough to allow continuous functions to be defined on it. In particular, a manifold does not have a metric defined. mani

A *chart* is a set of *functions* $\{x^1, \ldots, x^n\}$ which together map points on the manifold to $\mathbb{R}^n$. In other words, it is a *coordinate system*. The fact that the range of this map is (flat) $\mathbb{R}^n$ allows us to say that the manifold is *locally Euclidean*. chart coorc tem

Now consider a *path* on the manifold – this is just a continuous sequence of points. We distinguish this from a *curve*, $\lambda(t)$, say, which is a mapping from a parameter t to points on a path – two mappings which map to the same path but with different parametrisation are different curves. path curve

If we put these ideas together, and think of the functions $x^1(\lambda(t)), \ldots, x^n(\lambda(t))$, then we have a set of mappings from the curve parameter to the coordinates. The properties of the manifold tell us that these are smooth functions $x^i(t)$, so we can differentiate with respect to the parameter.

1.2 Defining the tangent vector

Now think of a function f which is defined on the manifold, and therefore at every point along the curve λ. The function f is a function of $\lambda(t)$, so it is also a function of the coordinates of the points along that curve, or

$$f = f\big(\lambda(t)\big) = f\Big(x^1\big(\lambda(t)\big), \ldots, x^n\big(\lambda(t)\big)\Big),$$

which we can write as just

$$f = f\big(x^1(t), \ldots, x^n(t)\big).$$

So how does f vary as we move along the curve? Easy:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\partial x^i}{\partial t} \frac{\partial f}{\partial x^i}.$$

However, since this is true of *any* function f, we can write instead

$$\frac{\mathrm{d}}{\mathrm{d}t} = \sum_i \frac{\partial x^i}{\partial t} \frac{\partial}{\partial x^i}. \tag{3.1}$$

Now consider the same path parameterised by $t_a = t/a$. We have

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t_a} &= \sum_i \frac{\partial x^i}{\partial t_a} \frac{\partial f}{\partial x^i} \\ &= a \sum_i \frac{\partial x^i}{\partial t} \frac{\partial f}{\partial x^i} \\ &= a \frac{\mathrm{d}f}{\mathrm{d}t} \end{aligned} \tag{3.2}$$

Consider another curve $\mu(s)$, which crosses curve $\lambda(t)$ at point P. We can therefore write, at P,

$$a\frac{\mathrm{d}f}{\mathrm{d}s} + b\frac{\mathrm{d}f}{\mathrm{d}t} = \sum_i \left(a\frac{\partial x^i}{\partial s} + b\frac{\partial x^i}{\partial t}\right)\frac{\partial f}{\partial x^i} = \sum_i \frac{\partial x^i}{\partial r}\frac{\partial f}{\partial x^i} = \frac{\mathrm{d}f}{\mathrm{d}r}, \tag{3.3}$$

for some further curve $\tau(r)$ which also passes through point P.

But now look what we have discovered. Whatever sort of thing $\mathrm{d}/\mathrm{d}t$ is, $a\mathrm{d}/\mathrm{d}t$ is the same type of thing (from Eq. (3.2)), and so is $a\mathrm{d}/\mathrm{d}s + b\mathrm{d}/\mathrm{d}t$. But we look at Sect. 1.1 of part 2, and realise that these derivative-things *at* P, which we'll write $(\mathrm{d}/\mathrm{d}t)_P$, satisfy the axioms of a *vector space*. Thus the things $(\mathrm{d}/\mathrm{d}t)_P$ are another example of things which can be regarded as vectors, or $\binom{1}{0}$ tensors. The thing $(\mathrm{d}/\mathrm{d}t)_P$ is referred to as a *tangent vector*. tange

A vector $\bar{V} = (\mathrm{d}/\mathrm{d}t)_P$ has rather a double life. Viewed as a derivative, $\bar{V}$ is just an operator which acts on a function f to give

$$\bar{V}f = \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)_P f = \left.\frac{\mathrm{d}f}{\mathrm{d}t}\right|_{t(P)},$$

the rate of change of f along the curve $\lambda(t)$, evaluated at P. There's nothing particularly exotic there. What we have just discovered, however, is that this object $(\mathrm{d}/\mathrm{d}t)_P$ can *also*, *separately*, be regarded as a vector in a vector space $T_P(M)$, and as such is a $\binom{1}{0}$ tensor, which is to say a thing which takes a *one-form* as an argument, to produce a number, $\langle\widetilde{\omega}, \bar{V}\rangle$, for some one-form $\widetilde{\omega}$ (we will see in a moment what this one-form is; it is not the function f). This dual aspect does seem confusing, and makes the object $\bar{V}$ seem more exotic than it really is, but it will (should be!) always clear from context which facet of the vector is being referred to at any point.

We'll refer to the set of these directional derivatives as $T_P(M)$, the *tangent plane of the manifold M at the point P*. It is very important to note that $T_P(M)$ and, say, $T_Q(M)$ – the tangent planes at two different points of the manifold – are *different spaces*, and have nothing tange to do with one another *a priori* (though this is ultimately why we introduce the connection in Sect. 2).

With this in mind, we can reread Eq. (3.1) as a vector expression, identifying the vectors

$$\bar{e}_i = \left(\frac{\partial}{\partial x^i}\right)_P \tag{3.4}$$

as a basis for the tangent-space, and the *numbers* $\partial x^i/\partial t$ as the components of the vector $\bar{V} = (\mathrm{d}/\mathrm{d}t)_P$ in this basis, or

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)_P &= \sum_i \frac{\partial x^i}{\partial t}\left(\frac{\partial}{\partial x^i}\right)_P \\ \bar{V} &= V^i \bar{e}_i .\end{aligned}$$

So, I've shown you that we *can* regard the $(\mathrm{d}/\mathrm{d}t)_P$ as vectors; the rest of this part of the course should convince you that this is additionally a *useful* thing to do.

1.3 The gradient one-form

[Schutz discusses these ideas in the opposite order, defining the gradient one-form before the tangent vector. See his §§3.3 and 5.2. Which do you think works better?]

Consider a function f, defined on the manifold. This is a *field*, which is to say it is a rule which associates an object – in this case the number which is the value of the function – with each point on the manifold (see Sect. 2.3 of part 2). Given this function, there is a particular one-form field which we can define (that is, a rule for associating a one-form with each point in the manifold), namely the *gradient one-form* $\tilde{\mathrm{d}} f$. Given a vector $\bar{V} = (\mathrm{d}/\mathrm{d}t)_P$, the tangent to a curve $\lambda(t)$, the gradient one-form is defined by its contraction with this vector: gradi form

$$\langle \tilde{\mathrm{d}} f, \bar{V} \rangle = \left\langle \tilde{\mathrm{d}} f, \frac{\mathrm{d}}{\mathrm{d}t} \right\rangle = \tilde{\mathrm{d}} f \left(\frac{\mathrm{d}}{\mathrm{d}t} \right) \equiv \left. \frac{\mathrm{d} f}{\mathrm{d}t} \right|_P . \tag{3.5}$$

The first two equalities here simply express notational equivalences; it is the third equality which is the definition of the gradient one-form's action. Ex.3

Now consider the gradient one-form associated with, not f, but one of the coordinate functions x^i (from Sect. 1.1, recall that the coordinates are just a set of functions on the manifold, and in this sense not importantly different from an arbitrary function f). We write these as simply $\tilde{\mathrm{d}} x^i$: what is their action on the basis vectors $\bar{e}_i = \partial/\partial x^i$ (from Eq. (3.4))? Directly from Eq. (3.5),

$$\tilde{\mathrm{d}} x^i \left(\frac{\partial}{\partial x^j} \right) = \frac{\partial x^i}{\partial x^j} = \delta^i{}_j , \tag{3.6}$$

so that, comparing this with Eq. (2.3a), we see that the set $\tilde{\omega}^i = \tilde{\mathrm{d}} x^i$ forms a basis for the one-forms which is *dual* to the vector basis $\bar{e}_i = \partial/\partial x^i$. dual

1.4 Basis transformations

What does a change of basis look like in this new notation? If we decide that we do not like the coordinate functions x^i and decide to use instead functions $x^{\bar{\imath}}$, how does this appear in our formalism, and how does it compare to Sect. 2.7 of part 2?

The new coordinates will generate a set of basis vectors

$$\bar{e}_{\bar{\imath}} = \frac{\partial}{\partial x^{\bar{\imath}}}. \tag{3.7}$$

This new basis will be related to the old one by a linear transformation

$$\bar{e}_{\bar{\imath}} = \Lambda^{j}_{\bar{\imath}} \bar{e}_j,$$

and the corresponding one-form basis will be related via the inverse transformation

$$\widetilde{\omega}^{\bar{\imath}} = \Lambda^{\bar{\imath}}_{j} \widetilde{\omega}^j$$

(recall example 2.7 of part 2). Thus, from Eq. (2.12),

$$\Lambda^{\bar{\imath}}_{j} = \widetilde{\omega}^{\bar{\imath}}(\bar{e}_j) = \widetilde{\mathrm{d}}x^{\bar{\imath}}\left(\frac{\partial}{\partial x^j}\right) = \frac{\partial x^{\bar{\imath}}}{\partial x^j} \tag{3.8a}$$

$$\Lambda^{j}_{\bar{\imath}} = \widetilde{\omega}^{j}(\bar{e}_{\bar{\imath}}) = \widetilde{\mathrm{d}}x^{j}\left(\frac{\partial}{\partial x^{\bar{\imath}}}\right) = \frac{\partial x^{j}}{\partial x^{\bar{\imath}}}. \tag{3.8b}$$

Note that the transformation matrix is defined as transforming the basis vectors and one-forms; as an immediate consequence it can transform the vector and one-form components also (as discussed in Sect. 2.7). Because of the above *choice* of basis vectors as the differentials of the coordinate functions (this is a *coordinate basis* – see Schutz for discussion of non-coordinate bases), the transformation matrix *also* describes a transformation between coordinate systems.

2 Covariant differentiation in flat spaces

We are now finally in a position to move on to the central tool of this part, the ideas of coordinate-independent differentiation of tensors, parallel transport, and curvature. We will make this move in two steps: first, we will learn how to handle the situation where the basis vectors of the space of interest are different at different points in the space, but confining ourselves to flat space, where we already know how to do most of the calculations; secondly, we will discover the rather simple step involved in transferring this knowledge to the case of fully curved spaces.

There are other ways of introducing the covariant derivative, which are very insightful, but more than a little abstract. Stewart [1, §1.7] introduces it in an axiomatic way which makes clear the tensorial nature of the covariant derivative from the very outset, as well as its linearities and some of its other properties. ‘Blue Schutz’ [2, ch.6] introduces it in a typically elegant way, *via* parallel transport, and emphasising the ultimate arbitrariness of the precise differentiation rule. Both of these routes define a connection which is more general than the one we are led to here, and only later specialise it to the *metric connection* which we are led to below. Chapter 10 of MTW [3] gives a very good, and visual, introduction to covariant differentiation, though approaching it from a somewhat different direction.

The point of this – the goal we are aiming for – is this: given some geometrical object $\bar{V}$ of physical interest (such as an electric field in a space, or a strain tensor in some medium), we want to know how it varies as we move around a space, in a way which doesn't depend on the coordinates we have chosen.

2.1 Differentiation of basis vectors

This section is to some extent another notation section, in that it is describing something you already know how to do, but in more elaborate and powerful language.

You will in the past have dealt with calculus in curvilinear coordinate systems and produced such results as the Laplacian in spherical polar coordinates being

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}.$$

We are now aiming for much the same destination, but by a slightly different route. This follows Schutz §§5.3–5.5 quite closely.

We will proceed by examining the basis vectors of (plane) polar coordinates, as expressed in terms of the cartesian basis vectors $\bar{e}_x$ and $\bar{e}_y$. In the next section we will see that our formalism is not restricted to this route.

The basis vectors of polar coordinates are

$$\bar{e}_r = \cos\theta\,\bar{e}_x + \sin\theta\,\bar{e}_y \tag{3.9a}$$

$$\bar{e}_\theta = -r\sin\theta\,\bar{e}_x + r\cos\theta\,\bar{e}_y. \tag{3.9b}$$

A little algebra shows that

$$\frac{\partial}{\partial r}\bar{e}_r = 0 \tag{3.10a}$$

$$\frac{\partial}{\partial\theta}\bar{e}_r = \frac{1}{r}\bar{e}_\theta \tag{3.10b}$$

$$\frac{\partial}{\partial r}\bar{e}_\theta = \frac{1}{r}\bar{e}_\theta \tag{3.10c}$$

$$\frac{\partial}{\partial\theta}\bar{e}_\theta = -r\bar{e}_r, \tag{3.10d}$$

so that we can see how the basis vectors *change* as we move to different points in the plane, unlike the cartesian basis vectors.

At any point in the plane, a vector $\bar{V}$ has components (V^r, V^θ) in the polar basis *at that point.* We can differentiate this vector with respect to, say, r, in the obvious way

$$\begin{aligned}\frac{\partial \bar{V}}{\partial r} &= \frac{\partial}{\partial r}(V^r \bar{e}_r + V^\theta \bar{e}_\theta) \\ &= \frac{\partial V^r}{\partial r}\bar{e}_r + V^r \frac{\partial \bar{e}_r}{\partial r} + \frac{\partial V^\theta}{\partial r}\bar{e}_\theta + V^\theta \frac{\partial \bar{e}_\theta}{\partial r},\end{aligned}$$

or, in index notation, with α running over the 'indexes' r and θ,

$$\begin{aligned}\frac{\partial \bar{V}}{\partial r} &= \frac{\partial}{\partial r}(V^\alpha \bar{e}_\alpha) \\ &= \frac{\partial V^\alpha}{\partial r}\bar{e}_\alpha + V^\alpha \frac{\partial \bar{e}_\alpha}{\partial r}.\end{aligned}$$

If, finally, we realise that there is nothing special about the coordinate r, and that we could write a similar expression involving x^β, which is *either* of the coordinates r and θ, we can write the perfectly general form (which we could in fact have written down directly),

$$\frac{\partial \bar{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\bar{e}_\alpha + V^\alpha \frac{\partial \bar{e}_\alpha}{\partial x^\beta}. \tag{3.11}$$

In cartesian coordinates, the second term in this expression is identically zero, since the basis vectors are the same everywhere on the plane, and so we can obtain the differential of a vector by simply differentiating its components (the first term above); this is not true when we are using curvilinear coordinates, and the second term comes in when we worry about how the basis vectors are different at different points on the plane.

Now, the second term above, $\partial\bar{e}_\alpha/\partial x^\beta$, is itself a vector, so that it is a linear combination of the basis vectors, with coefficients $\Gamma^\mu_{\alpha\beta}$:

$$\frac{\partial\bar{e}_\alpha}{\partial x^\beta} = \Gamma^\mu_{\alpha\beta}\bar{e}_\mu. \tag{3.12}$$

This symbol $\Gamma^\mu_{\alpha\beta}$ is called the *Christoffel symbol*, and the $n \times n \times n$ numbers which it represents encode all the information we need about how the coordinates, and their associated basis vectors, change within the space. It is not a tensor – it is merely a collection of numbers – so its indexes are not staggered (just like the transformation matrix Λ).

We have done all the work to calculate the Christoffel symbol(s) for polar coordinates. If we compare Eq. (3.10) with Eq. (3.12), we see

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = 1/r, \qquad \Gamma^r_{\theta\theta} = -r, \tag{3.13}$$

with all other components zero.

2.2 The covariant derivative in flat spaces

Importantly, there is nothing in the definition of the Christoffel symbols, Eq. (3.12), or their calculated values, Eq. (3.13), which refers to the cartesian basis which we (incidentally) used when working them out in Eq. (3.10). That is, we are at this point free of any dependence on a particular coordinate system.

Notation: If we rewrite Eq. (3.11) including Eq. (3.12), relabel and reorder, we find

$$\frac{\partial \bar{V}}{\partial x^\beta} = \left(\frac{\partial V^\alpha}{\partial x^\beta} + V^\mu \Gamma^\alpha_{\mu\beta} \right) \bar{e}_\alpha. \tag{3.14}$$

For each β this is a vector at each point in the space – that is to say, it is a vector field – with components given by the term in brackets. We denote these components of the vector field by the notation $V^\alpha{}_{;\beta}$, with the semicolon denoting covariant differentiation. If we further denote the derivative of the component $\partial V^\alpha / \partial x^\beta = V^\alpha{}_{,\beta}$, then we can write

$$\frac{\partial \bar{V}}{\partial x^\beta} = V^\alpha{}_{;\beta} \bar{e}_\alpha \tag{3.15a}$$

$$V^\alpha{}_{;\beta} = V^\alpha{}_{,\beta} + V^\mu \Gamma^\alpha_{\mu\beta}. \tag{3.15b}$$

It is important to be clear about what you are looking at, here. The objects $V^\alpha{}_{;\beta}$ are numbers which are the *components*, indexed by α, of a *set of vectors*, indexed by β. They look rather like tensor components, however, and we are about to deduce that that is exactly what they are in fact. But components of which tensor? Ex.3.

Final step: Look back at Eq. (3.7), and notice that the differential $\partial/\partial x^\beta$ in Eq. (3.14) is Ex.3. associated with the basis vector $\bar{e}_\beta$. Indeed, the (vector) differential $\partial/\partial x^\beta$ is a *linear* function Ex.3. of the vector $\bar{e}_\beta$ (if you doubled the length of $\bar{e}_\beta$, you would halve the values of x^β, and thus double $\partial/\partial x^\beta$). Another way of saying that is that there exists a $\binom{1}{1}$ tensor, which we shall call $\nabla \bar{V}$, which we shall define by saying that the action of it on the vector $\bar{e}_\beta$ is the vector $\partial \bar{V}/\partial x^\beta$ in Eq. (3.14). Using the notation of part 2, we could write

$$(\nabla \bar{V})(\tilde{\cdot}; \bar{e}_\beta) \equiv \frac{\partial \bar{V}}{\partial x^\beta}(\tilde{\cdot}) \tag{3.16}$$

as the *definition* of the tensor $\nabla \bar{V}$. For notational convenience, we prefer to write this as

$$\nabla_{\bar{e}_\beta} \bar{V} = \frac{\partial \bar{V}}{\partial x^\beta}. \tag{3.17}$$

This tensor $\nabla \bar{V}$ is called the *covariant derivative of* $\bar{V}$, and its components are

$$(\nabla \bar{V})^\alpha{}_\beta \equiv (\nabla_{\bar{e}_\beta} \bar{V})^\alpha \equiv (\nabla_\beta \bar{V})^\alpha = V^\alpha{}_{;\beta}, \tag{3.18}$$

where the first equivalence is the definition of the tensor, restated from the text above Eq. (3.16), the second is a notational convenience, which applies in the case where the argument vector is a basis vector, and the equality connects the others to the definition Eq. (3.15) or Eq. (3.17).

Here's where we've got to: we've managed to define a *tensor* field related to $\bar{V}$, called the covariant derivative, and written $\nabla \bar{V}$, which (since it is a tensor) is independent of *any* coordinate system, and so doesn't depend on any coordinate system, and doesn't pick out any coordinate system as special. If we need its components in a particular system $\{x^\mu\}$, however, because we need to do some calculations, we can find them easily, *via* Eq. (3.15), or by transforming the components from a system where we already know them (such as cartesian coordinate) into the system $\{x^\mu\}$ – we know we can do this because we know that $\nabla \bar{V}$ is a tensor, so we know how its components transform.

Finally, here, note that a scalar is independent of any coordinate system, therefore all the complications of this section, which essentially involve dealing with the fact that basis vectors are different at different points on the manifold, disappear, and the covariant derivative of a scalar is simply the partial derivative:

$$\nabla_\beta f = \frac{\partial f}{\partial x^\beta}. \tag{3.19}$$

Comparing this with Eq. (3.5) (with $\bar{e}_\beta = \partial/\partial x^\beta$ instead of $\bar{V}$), we see that we can identify the covariant derivative of a function with the gradient one-form:

$$\nabla f = \widetilde{\mathrm{d}} f.$$

From this we can deduce the expression for the covariant derivative of a one-form, which we shall simply quote as:

$$(\nabla_\beta \widetilde{p})_\alpha \equiv (\nabla \widetilde{p})_{\alpha\beta} \equiv p_{\alpha;\beta} = p_{\alpha,\beta} - p_\mu \Gamma^\mu_{\alpha\beta}, \tag{3.20}$$

(note the sign difference from Eq. (3.15)). The derivative of a $\binom{1}{1}$ tensor is

$$\nabla_\beta T^\mu{}_\nu \equiv T^\mu{}_{\nu;\beta} = T^\mu{}_{\nu,\beta} + \Gamma^\mu_{\alpha\beta} T^\alpha{}_\nu - \Gamma^\alpha_{\nu\beta} T^\mu{}_\alpha. \tag{3.21}$$

Note how systematic this expression is, and that it is systematically extensible to higher orders of tensor – there is one $+\Gamma$ term for each upper tensor index, and one $-\Gamma$ term for each lower index. The expression looks hard to remember, but is easier than it looks.

Also, the product rule applies

$$\nabla_\beta (p_\mu V^\mu) = p_{\mu;\beta} V^\mu + p_\mu V^\mu{}_{;\beta}. \tag{3.22}$$

See Schutz §5.3 for details.

2.3 The metric and the Christoffel symbol

The covariant derivative, and the Christoffel symbol, give us information about how, and how quickly, the basis vectors change as we move about a space. It is therefore no surprise to find that there is a deep connection between these and the *metric*, which gives information about distances within a space.

Remember that the metric (a $\binom{0}{2}$ tensor) allows us to identify a particular one-form associated with a given vector:

$$\widetilde{V} = \mathsf{g}(\bar{V}, \cdot).$$

Note that this is a purely geometrical (ie, coordinate-independent) equation. Recall that in cartesian coordinates (a) the components of a vector and its associated one-form are equal, and (b) the basis vectors are constant, so that covariant differentiation is just straight partial differentiation of the components, so that (c) in cartesian coordinates the components of the covariant derivative (with respect to a basis vector $\bar{e}_\beta$) of a vector and of its associated one-form are equal:

$$\nabla_\beta \widetilde{V} = \mathsf{g}(\nabla_\beta \bar{V}, \cdot).$$

But this is *also* a purely geometrical equation; so that, even though we justified it using a particular (cartesian) coordinate system, it must be true in *all* coordinate systems.

In components (and in all coordinate systems),

$$V_\alpha = g_{\alpha\mu} V^\mu \tag{3.23}$$

$$V_{\alpha;\beta} = g_{\alpha\mu} V^\mu{}_{;\beta}. \tag{3.24}$$

Note that the latter equation is *not* trivial. From the properties of the metric we know that there exists *some* tensor which has components $A_{\alpha\beta} = g_{\alpha\mu} V^\mu{}_{;\beta}$: what this expression tells us is the non-trivial statement that this $A_{\alpha\beta}$ is exactly $V_{\alpha;\beta}$.

That is to say that we did *not* get Eq. (3.24) by differentiating Eq. (3.23), though it looks rather as if we did. What *do* we get by differentiating Eq. (3.23)? By the product rule Eq. (3.22)

$$\bar{V}_{\alpha;\beta} = g_{\alpha\mu;\beta} V^{\mu} + g_{\alpha\mu} V^{\mu}{}_{;\beta}.$$

But comparing this with Eq. (3.24), we see that the first term on the right-hand side must be zero, for arbitrary $\bar{V}$. Thus, in all coordinate systems (and relabelling)

$$g_{\alpha\beta;\mu} = 0. \tag{3.25}$$

We have not exhausted the link between covariant differentiation and the metric. The two are related *via*

$$\Gamma^{\gamma}_{\beta\mu} = \tfrac{1}{2} g^{\alpha\gamma} (g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha}). \tag{3.26}$$

The proof is in Schutz §5.4, leading up to his equation (5.75); it is not hard to follow, just a little tedious. It depends on first proving that

$$\Gamma^{\mu}_{\alpha\beta} = \Gamma^{\mu}_{\beta\alpha}, \qquad \text{in all coordinate systems.} \tag{3.27}$$

Equation (3.26) completely cuts the link between the Christoffel symbol and cartesian coordinates, which might have lingered in your mind after Sect. 2.2 – once we have a metric, we can work out the Christoffel symbol's components immediately. Ex.3

Ex.3

3 The covariant derivative in curved spaces

Having done all this work to develop covariant differentiation in flat space, but in purely geometrical terms, it might be a surprise to discover that there is actually rather little to do to bring this over to the most general case of curved spaces. See Schutz §§6.2–4.

The first step is to define carefully the notion of a *local inertial frame*.

3.1 Local inertial frames – the local flatness theorem

Recall from Sect. 1.1 that a manifold is little more than a collection of points. What gives this manifold shape is the metric tensor g, which is a symmetric $\binom{0}{2}$ tensor which, in a particular coordinate system, has the components $g_{\alpha\beta}$, which we can choose more-or-less how we like. In a different coordinate system, this same tensor will have different components $g_{\bar{\alpha}\bar{\beta}}$. The question is, can we find a coordinate system in which the metric has the particular form $\eta_{\bar{\alpha}\bar{\beta}} = \mathrm{diag}(-1, 1, 1, 1)$? That is, can we find a coordinate transformation $\Lambda^{\bar{\alpha}}_{\alpha}$ which transforms the coordinates x^{α} into the coordinates $x^{\bar{\alpha}}$ in which the metric is diagonal?

If the *matrix* $g_{\alpha\beta}$ does not have three positive and one negative eigenvalues (ie, a *signature* of $+3-1=+2$) then no, we cannot, and the metric in question is uninteresting to us because it cannot describe our universe. If the metric does have a signature of $+2$, however, then it is a theorem of linear algebra that we can indeed find a transformation to coordinates in which the metric is diagonal *at a point*.

But we can do better than this. Recall that both $g_{\alpha\beta}$ and $\Lambda^{\bar{\alpha}}_{\alpha}$ are continuous functions of position; within the constraints that g be symmetric and Λ be invertible, they are arbitrary. By choosing the numbers $\Lambda^{\bar{\alpha}}_{\alpha}$ *and* their first derivatives, we can find coordinates which have their origin at P and in which

$$g_{\bar{\alpha}\bar{\beta}}(x^{\bar{\mu}}) = \eta_{\bar{\alpha}\bar{\beta}} + O((x^{\bar{\mu}})^2),$$

(compare Taylor's theorem) or

$$g_{\bar{\alpha}\bar{\beta}}(P) = \eta_{\bar{\alpha}\bar{\beta}} \tag{3.28a}$$

$$g_{\bar{\alpha}\bar{\beta},\bar{\gamma}}(P) = 0 \tag{3.28b}$$

$$g_{\bar{\alpha}\bar{\beta},\bar{\gamma}\bar{\delta}}(P) \neq 0. \tag{3.28c}$$

This is the *local flatness theorem*, and the coordinates $x^{\bar{\mu}}$ represent a *local inertial frame*, or LIF.

local
theor
local
frame

These coordinates are also known as ‘normal’ or ‘geodesic’ coordinates, and geodesics expressed in these coordinates have a particularly simple form. Also, in these coordinates, it turns out that $\Gamma^i_{jk} = 0$ at a point, which is just another way of saying that this space is locally flat.

Schutz’s proof of the theorem at the end of his §6.2 is very illuminating.

3.2 Covariant derivatives in curved spaces

You know how to differentiate things. For some function f,

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{3.29}$$

That's straightforward because it's obvious what $f(x+h) - f(x)$ means. Surely we can do a similar thing with vectors on a manifold. Not trivially, because remember that the vectors at P are not defined on the *manifold* but on the *tangent plane* $T_P(M)$ at a point P, and so the vectors at a different point Q are in a completely different space $T_Q(M)$, so it's not obvious how to 'subtract' one vector from the other. Differentiation on the manifold consists of finding ways to define just that 'subtraction'.

There are several ways to do this. One produces the 'Lie derivative', which is important in many respects, but which we will not examine.

The Lie derivative is a coordinate-independent derivative defined in terms of a vector field. A vector field $\bar{X}$ has *integral curves* such that, at each point p on the integral curve, the curve's tangent vector is $\bar{X}(p)$. As an example, stream lines in a fluid are integral curves of the fluid's velocity vector field.

The Lie derivative of a function at a point p, written $(\pounds_{\bar{X}} f)_p$, is defined as the rate of change of the function along the (unique) integral curve of $\bar{X}$ going through p, and Lie derivatives of higher-order tensors are defined in an analogous way. The disadvantage of this type of derivative is that it clearly depends on an auxiliary vector field $\bar{X}$; but the compensating advantage is that it does *not* depend on a metric tensor, or any other definition of distance. These make it less useful than the covariant derivative for most GR applications, but it remains useful in other contexts, such as those where there is already an important vector field present, including applications in fluid dynamics. For details, see [2] or [1]; or look at exercise 39 in Schutz's §6.9.

The other way to define this 'subtraction' uses the notion of 'parallel transport', which we define and examine now.

You *parallel transport* a vector along a curve if the vectors at any two infinitesimally paral port separated points are deemed parallel, in the sense of having the same length and pointing in the same direction (a diagram helps here). The rule for deciding whether two such vectors are parallel isn't specified here, and is broadly up to you, but we'll come back to that.

This gives us a way of talking about 'subtraction'. Take a vector field $\bar{V}$ on the manifold, and two points P and Q which are both on some curve $\lambda(t)$, with tangent vector $\bar{U}$. We can take the vector $\bar{V}(Q)$ at Q and parallel transport it back to P; at that point it is in the same space $T_P(M)$ as the vector $\bar{V}(P)$ so we can unambiguously subtract them. These two points are a parameter distance $t(Q) - t(P)$ apart, so we can divide by that distance, and find the limit as that distance goes to zero, and thus reconstruct all the components we need to define a differential just like Eq. (3.29). The differential we get by this process is the *covariant derivative of* $\bar{V}$ *along* $\bar{U}$, written $\nabla_{\bar{U}}\bar{V}$.

The covariant derivative depends on using parallel transport as a way of connecting vectors in two different tangent planes. The covariant derivative is sometimes also called the *connection*, and the Christoffel symbols the *connection coefficients*.

If $\bar{V}(Q)$ starts off as just the parallel-transported version of $\bar{V}(P)$, then when we parallel-transport it back to P we'll get just $\bar{V}(P)$ again, so that this covariant derivative will be zero; thus

$$\nabla_{\bar{U}}\bar{V} = 0 \Leftrightarrow (\bar{V} \text{ is parallel transported along } \bar{U}). \tag{3.30}$$

The crucial thing here is that nowhere in this account of the covariant derivative have we mentioned coordinates at all.

Writing down the details of the construction of this derivative would be notationally intricate and take us a little too far afield. If you want details, they're in blue-Schutz or Stewart. Also, the definition of parallelism via the LIF, below, is not the only one possible, but picks out a particular derivative and set of connection coefficients, called the 'metric connection'. Only with this connection are Eq. (3.31) and Eq. (3.32) true.

We've actually said rather little, here, because although this has, I hope, made clear how closely linked are the ideas of the covariant derivative and parallel transport, we haven't said anything about exactly how we go about choosing a definition of parallelism, and we haven't seen how this links to the covariant derivative we introduced in Sect. 2. The link is the locally-flat local inertial frame. Although the general idea of parallel transport, and in particular the definition we introduce below, may seem obvious or intuitive, do remember that there is a good deal or arbitrariness in its actual definition.

Consider the coordinates representing the LIF at the point P. This is a flat cartesian space (but not euclidean, remember, since it does not have a euclidean metric). That means that the basis vectors are constant – their derivatives are zero. A definition of parallelism now jumps out at us: *two nearby vectors are parallel if their components in the LIF are the same.* But this is the definition of parallelism that was implicit in the differentiations we used in sections 2.1 and 2.2, leading up to Eq. (3.11), and so the covariant derivative we end up with is the same one: the tensor $\nabla\bar{V}$ as defined in this section is the same as the tensor $\nabla\bar{V}$ of Eq. (3.18). In other words, in this cartesian frame, covariant differentiation *as defined in this section* is the same as ordinary differentiation, and

$$V^{\alpha}{}_{;\beta} = V^{\alpha}{}_{,\beta} \qquad \text{in LIF.}$$

Now, this is true for any tensor, so specifically

$$g_{\alpha\beta;\mu} = g_{\alpha\beta,\mu} = 0 \qquad \text{at } P,$$

by Eq. (3.28a). But this is a tensor equation, so it is true in any coordinate system, and since there is nothing special about the point P, it is true at all points of the manifold:

$$g_{\alpha\beta;\mu} = 0 \qquad \text{in any coordinate system.} \tag{3.31}$$

As mentioned at the end of Sect. 2.3, from Eq. (3.31) we can, if $\Gamma^{\mu}_{\alpha\beta} = \Gamma^{\mu}_{\beta\alpha}$ (which is the case for this definition of covariant differentiation), deduce that

$$\Gamma^{\alpha}_{\mu\nu} = \tfrac{1}{2} g^{\alpha\beta}(g_{\beta\mu,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta}). \tag{3.32}$$

See also Schutz's discussion of geodesics on his pp.166–7, which elaborates the idea of parallelism introduced here.

4 Geodesics

Consider a curve $\lambda(t)$ and its tangent vectors $\bar{U}$ (that is, the set of vectors $\bar{U}$ is a field which is defined at least at all the points along the curve λ). If we have another vector field $\bar{V}$, then the vector $\nabla_{\bar{U}}\bar{V}$ tells us how much $\bar{V}$ changes as we move along the curve λ which $\bar{U}$ is the tangent to. What happens if, instead of the arbitrary vector field $\bar{V}$ we take the covariant derivative of $\bar{U}$ itself? In general, $\nabla_{\bar{U}}\bar{U}$ will not be zero – if the curve 'turns a corner', then the tangent vector after the corner will no longer be parallel to the tangent before the corner. The meaning of 'parallel' here is exactly the same as the meaning of 'parallel' which was built in to the definition of the covariant derivative in the passage after Eq. (3.30). Curves which do not do this – that is, curves such that all the tangent vectors are parallel to each other – are the nearest thing to a straight line in the space, and indeed *are* a straight line in a flat space. A curve such as this is called a *geodesic*. Thus the definition of a geodesic is:

$$\nabla_{\bar{U}}\bar{U} = 0. \Leftrightarrow (\bar{U} \text{ is the tangent to a geodesic}) \tag{3.33}$$

Equation (3.33) has a certain spartan elegance, but if we are to do any calculations to discover what the path of the geodesic actually is, we need to unpack it.

The object $\nabla_{(\cdot)}\bar{U}$ is a $\binom{1}{1}$ tensor, as you will recall, with its vector argument denoted by the $(\cdot)$ (peculiar notation, I know…). Since it is a tensor, it is linear in this argument. That is, for any vector $\bar{A}$ and scalar a, $\nabla_{a\bar{A}}\bar{U} = a\nabla_{\bar{A}}\bar{U}$, and specifically $\nabla_{A^\mu \bar{e}_\mu}\bar{U} = A^\mu \nabla_{\bar{e}_\mu}\bar{U} \equiv A^\mu \nabla_\mu \bar{U}$. The vector $\bar{U}$ has components

$$\bar{U} = U^\beta \bar{e}_\beta,$$

and so Eq. (3.33) can be written

$$U^\beta \nabla_\beta \bar{U} = U^\beta U^\alpha{}_{;\beta}\bar{e}_\alpha = 0$$

(recalling Eq. (3.18)). The α-component of this equation is, using Eq. (3.15b),

$$U^\beta U^\alpha{}_{;\beta} = U^\beta U^\alpha{}_{,\beta} + U^\beta U^\mu \Gamma^\alpha_{\beta\mu} = 0.$$

Let t be the parameter along the geodesic (that is, there is a parameterisation of the geodesic, $\lambda(t)$, with parameter t, which $\bar{U}$ is the tangent to). Then $U^\beta = \mathrm{d}x^\beta/\mathrm{d}t$ and $U^\alpha{}_{,\beta} = \partial/\partial x^\beta (\mathrm{d}x^\alpha/\mathrm{d}t)$, and pretty immediately we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}t}\right) + \Gamma^\alpha_{\beta\mu}\frac{\mathrm{d}x^\beta}{\mathrm{d}t}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = 0. \qquad (3.34)$$

This is the *geodesic deviation equation*. For each α it is a second order differential equation with initial conditions comprising the initial position $x_0^\alpha = x^\alpha(t_P)$ (if the parameter t has value t_P at point P) and initial direction/speed $U_0^\alpha = \mathrm{d}x^\alpha/\mathrm{d}t|_{t_P}$. The theory of differential equations tells us that this equation does have a unique solution.

geod
tion

A parameter t for which we can write down the geodesic equation Eq. (3.34) is termed an *affine parameter*, and if t is an affine parameter, it is easy to confirm that $\phi = at + b$, where a and b are constants, is an affine parameter also.

affin
ter

An affine parameter is one which, in MTW's words [3, §1.5], is 'defined so that motion looks simple'. You can reasonably measure time in seconds since midnight, or minutes (seconds/60), or minutes since noon (seconds/60 − 720). These are all affine transformations, and they share the property that unaccelerated motion is a linear function of time. If you were reckless enough to measure time in units of seconds-squared, then unaccelerated (that is, simple) motion would look very complicated indeed. Another way of saying this is that *an affine parameter is the time coordinate of some inertial system*, and all that means it that an affine parameter is the time shown on some free-falling 'good' clock. There are further remarks about affine parameters in Sect. 4.1.

Ex.3

The connection (or rather the class of connections) we have defined here (see Sect. 3.2) is constructed in such a way as to preserve parallelism. Such a connection is an *affine connection* – the word 'affine' comes from a Latin root meaning 'neighbouring'. Other types of connection are possible; see blue-Schutz [2, §6.14] if you're *very* keen.

A geodesic is a curve of extremal length. In a space with a metric with the signature of

GR, it is a curve of maximal length; in a euclidean space it is a curve of minimal length: for Euclid, a straight line is the shortest distance between two points.

Note on metric connections (extremely non-examinable): in other of these asides I have emphasised that this metric connection is not the only one definable. Since geodesics are defined in terms of the connection, it does indeed follow that the geodesics implied by these other connections are different from the geodesics of the metric connection, and specifically are not the curves of extremal length. This is bound up with the property Eq. (3.31), and the observation that only with the metric connection is the dot product $\mathsf{g}(\bar{A}, \bar{B})$ invariant under parallel transport. This is one reason why the metric connection is so important, to the point of being essentially ubiquitous in general relativity.

4.1 The variational principle and the geodesic equation

We can prove directly that the geodesic is a curve of extremal length, by deriving the geodesic equation explicitly from a variational principle.

For a given curve through spacetime, parameterised by λ, the length of the curve is given by

$$l = \int_{\text{curve}} \mathrm{d}s = \int_{\text{curve}} \left| g_{\alpha\beta} \mathrm{d}x^\alpha \mathrm{d}x^\beta \right|^{1/2} = \int_{\lambda_0}^{\lambda_1} \left| g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta \right|^{1/2} \mathrm{d}\lambda, \equiv \int_{\lambda_0}^{\lambda_1} \dot{s}\, \mathrm{d}\lambda,$$

where

$$\dot{s} = \left| g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta \right|^{1/2}$$

expresses the relationship between parameter distance and proper distance, and where dots indicate $\mathrm{d}/\mathrm{d}\lambda$. We wish to find a curve which is extremal, in the sense that its length l is unchanged under first-order variations in the curve, for fixed λ_0 and λ_1. The *calculus of variations* (which as physicists you are most likely to have met in the context of classical mechanics) tells us that such an extremal curve $x^\mu(\lambda)$ is the solution of the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial \dot{s}}{\partial \dot{x}^\mu}\right) - \frac{\partial \dot{s}}{\partial x^\mu} = 0.$$

Have a go, yourself, at deriving the geodesic equation from this, before reading the discussion below (at an appropriate point, you will need to choose $\ddot{s} = 0$).

For $\dot{s}$ as given above, we find fairly directly that

$$-\frac{1}{2}\frac{\ddot{s}}{\dot{s}^2} 2 g_{\mu\beta}\dot{x}^\beta + \frac{1}{2\dot{s}}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(2 g_{\mu\beta}\dot{x}^\beta\right) - \frac{1}{2\dot{s}} g_{\alpha\beta,\mu}\dot{x}^\alpha \dot{x}^\beta = 0. \qquad (3.35)$$

To simplify this, we can *choose* at this point to restrict ourselves to parameterisations of the curve which are such that $\mathrm{d}s/\mathrm{d}\lambda$ is constant along the curve, so that $\ddot{s} = 0$; this λ is an affine parameter as described above. With this choice, and multiplying overall by $\dot{s}$, we find

$$g_{\mu\beta,\nu}\dot{x}^{\beta}\dot{x}^{\nu} + g_{\mu\beta}\ddot{x}^{\beta} - \tfrac{1}{2}g_{\alpha\beta,\mu}\dot{x}^{\alpha}\dot{x}^{\beta} = 0$$

which, after relabelling and contracting with $g^{\mu\nu}$, and comparing with Eq. (3.32), reduces to

$$\ddot{x}^{\mu} + \Gamma^{\mu}_{\alpha\beta}\dot{x}^{\alpha}\dot{x}^{\beta} = 0, \tag{3.36}$$

the geodesic equation of Eq. (3.34).

As well as showing the direct connection between the geodesic equation and this deep variational principle, and thus making clear the idea that a geodesic is a 'shortest distance', this also confirms the significance of affine parameters which was touched on in Sect. 4. There is a 'geodesic equation' for non-affine parameters (namely Eq. (3.35)), but only when we choose an affine parameter λ, does this equation take the relatively simple form of Eq. (3.34) or Eq. (3.36). The general solution of Eq. (3.35) is the same *path* as the geodesic, but because of the non-affine parameterisation it is not the same *curve*, and is not, formally, a geodesic.

Schutz discusses this at the end of his §6.4, and the exercises corresponding to it.

5 Curvature

We now come, finally, to the coordinate-independent description of *curvature*. We approach it through the idea of parallel transport, as described in Sect. 3.2, and specifically through the idea of transporting a vector round a closed path. This section follows Schutz §6.5. MTW [3] chapter 11 is good on this. curva

5.1 The Riemann tensor

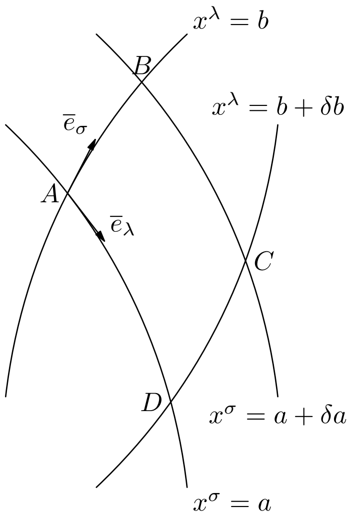

Figure 1

Consider the path following lines of constant coordinate, in an arbitrary coordinate system. Figure 1 shows a loop in the plane of two coordinates x^σ and x^λ The line joining A and B, and the line from D to C, have coordinate x^σ varying along a line of constant x^λ, and lines B–C and A–D have x^λ varying along a line of constant x^σ. We have a vector $\bar{V}$ at A, which we parallel-transport to B, C, D and back to A, and we want to find out how different the vector is after its circuit from how it was when it started.

Parallel-transporting the vector from A to B involves transporting $\bar{V}$ along the vector field $\bar{e}_\sigma$. From Eq. (3.30), this means that $\nabla_\sigma \bar{V} = 0$, or $V^\alpha{}_{;\sigma} = 0$. That is (from Eq. (3.15b)),

$$\frac{\partial V^\alpha}{\partial x^\sigma} = V^\alpha{}_{,\sigma} = -\Gamma^\alpha_{\mu\sigma} V^\mu. \tag{3.37}$$

Now, the components of the vector at B are

$$\begin{aligned} V^\alpha(B) &= V^\alpha(A) + \int_A^B \frac{\partial V^\alpha}{\partial x^\sigma} \mathrm{d}x^\sigma \\ &= V^\alpha(A) - \int_A^B \Gamma^\alpha_{\mu\sigma} V^\mu \mathrm{d}x^\sigma \\ &= V^\alpha(A) - \int_{x^\sigma = a}^{a+\delta a} \Gamma^\alpha_{\mu\sigma} V^\mu \Big|_{x^\lambda = b} \mathrm{d}x^\sigma, \end{aligned}$$

where the integrand is evaluated along the line $\{x^\lambda = b\}$ from $x^\sigma = a$ to $x^\sigma = a + \delta a$. Doing the same thing for the other sides of the curve, we find:

$$\begin{aligned} \delta V^\alpha = V^\alpha(A_{\text{final}}) - V^\alpha(A_{\text{init}}) = &- \int_{x^\sigma = a}^{a+\delta a} \Gamma^\alpha_{\beta\sigma} V^\beta \Big|_{x^\lambda = b} \mathrm{d}x^\sigma \\ &- \int_{x^\lambda = b}^{b+\delta b} \Gamma^\alpha_{\beta\lambda} V^\beta \Big|_{x^\sigma = a+\delta a} \mathrm{d}x^\lambda \\ &+ \int_{x^\sigma = a}^{a+\delta a} \Gamma^\alpha_{\beta\sigma} V^\beta \Big|_{x^\lambda = b+\delta b} \mathrm{d}x^\sigma \\ &+ \int_{x^\lambda = b}^{b+\delta b} \Gamma^\alpha_{\beta\lambda} V^\beta \Big|_{x^\sigma = a} \mathrm{d}x^\lambda. \end{aligned} \tag{3.38}$$

At this point we can take advantage of the fact that δa and δb are small, by construction, ignore terms in δa^2 and δb^2, and thus take the integrands to be constant along the interval of integration (by expanding the integrand in a taylor series, convince yourself that $\int_a^{a+\delta a} f(x) \mathrm{d}x =$

$\delta a f(a) + O(\delta a^2)$). We don't know what the $\Gamma^\alpha_{\beta\lambda} V^\beta|_{x^\sigma=a+\delta a}$ and $\Gamma^\alpha_{\beta\sigma} V^\beta|_{x^\lambda=b+\delta b}$ are (of course, since we are doing this calculation for perfectly general Γ), but since δa is small, we *can* estimate them using Taylor's theorem, finding

$$\Gamma^\alpha_{\beta\lambda} V^\beta\Big|_{x^\sigma=a+\delta a} = \cdots\Big|_{x^\sigma=a} + \delta a \frac{\partial}{\partial x^\sigma} \Gamma^\alpha_{\beta\lambda} V^\beta\Bigg|_{x^\sigma=a} + O(\delta a^2).$$

Inserting this, and the similar expression involving δb, into Eq. (3.38), and ignoring terms of $O(\delta a^2, \delta b^2)$, we have

$$\begin{aligned}\delta V^\alpha \approx &+ \int_{x^\sigma=a}^{a+\delta a} \delta b \ \frac{\partial}{\partial x^\lambda} \Gamma^\alpha_{\beta\sigma} V^\beta\Bigg|_{x^\sigma=a, x^\lambda=b} \mathrm{d}x^\sigma \\ &- \int_{x^\lambda=b}^{b+\delta b} \delta a \ \frac{\partial}{\partial x^\sigma} \Gamma^\alpha_{\beta\lambda} V^\beta\Bigg|_{x^\sigma=a, x^\lambda=b} \mathrm{d}x^\lambda.\end{aligned}$$

However, the integrands here are now constant with respect to the variable of integration, so the integrals are easy:

$$\delta V^\alpha \approx \delta a \delta b \left[\frac{\partial}{\partial x^\lambda} \left(\Gamma^\alpha_{\beta\sigma} V^\beta \right) - \frac{\partial}{\partial x^\sigma} \left(\Gamma^\alpha_{\beta\lambda} V^\beta \right) \right],$$

with all quantities evaluatated at the point A. If we now use Eq. (3.37) to get rid of the differentials of V^β, we find, to first order

$$\delta V^\alpha = \delta x^\sigma \delta x^\lambda [\Gamma^\alpha_{\beta\sigma,\lambda} - \Gamma^\alpha_{\beta\lambda,\sigma} - \Gamma^\alpha_{\mu\sigma} \Gamma^\mu_{\beta\lambda} + \Gamma^\alpha_{\mu\lambda} \Gamma^\mu_{\beta\sigma}] V^\beta, \tag{3.39}$$

where we have written δa and δb as δx^σ and δx^λ respectively.

Let us examine this result. The left-hand side is the α component of a vector $\delta \bar{V}$ (we know this is a vector since it is the difference of two vectors located at the same point A; recall the vector-space axioms); we obtain that component δV^α by acting on the vector $\delta \bar{V}$

with the basis one-form $\widetilde{\omega}^{\alpha}$. The right-hand side clearly depends on the vector $\bar{V}$ (also at the point A), whose components are V^{β}. The construction in Fig. 1, which crucially has the area enclosed by constant-coordinate lines, depends on multiples of the basis vectors, $\delta a \bar{e}_{\sigma}$ and $\delta b \bar{e}_{\lambda}$. We can see that the *number* δV^{α} depends *linearly* on each of these four objects – one one-form and three vectors. This leads us to identify the numbers within the square brackets of Eq. (3.39) as the components of a $\binom{1}{3}$ tensor

$$R^{\alpha}{}_{\beta\mu\nu} = \Gamma^{\alpha}_{\beta\nu,\mu} - \Gamma^{\alpha}_{\beta\mu,\nu} + \Gamma^{\alpha}_{\sigma\mu}\Gamma^{\sigma}_{\beta\nu} - \Gamma^{\alpha}_{\sigma\nu}\Gamma^{\sigma}_{\beta\mu}, \tag{3.40}$$

(after some relabelling) called the *Riemann curvature tensor* (this notation is consistent with Schutz; numerous other conventions exist – see the discussion in part 1). Thus Eq. (3.39) becomes

Riem
ture t

$$\delta V^{\alpha} = R^{\alpha}{}_{\beta\sigma\lambda} V^{\beta} \delta x^{\sigma} \delta x^{\lambda}. \tag{3.41}$$

This tensor tells us how the vector $\bar{V}$ varies after it is parallel-transported on an arbitrary excursion in the area local to point A (that is, for small δa and δb); that is, it encodes all the information about the local shape of the manifold.

Another way to see the significance of the Riemann tensor is to consider the effect of taking the covariant derivative of a vector with respect to first one then another of the coordinates, $\nabla_{\alpha}\nabla_{\beta}\bar{V}$. Defining the *commutator*

comm

$$[\nabla_{\alpha}, \nabla_{\beta}]V^{\mu} \equiv \nabla_{\alpha}\nabla_{\beta}V^{\mu} - \nabla_{\beta}\nabla_{\alpha}V^{\mu}, \tag{3.42}$$

we find that

$$[\nabla_{\alpha}, \nabla_{\beta}]V^{\mu} = R^{\mu}{}_{\nu\alpha\beta}V^{\nu}. \tag{3.43}$$

This, or something like it, might not be a surprise. We discovered the Riemann tensor by taking a vector for a walk round the circuit $ABCDA$ in Fig. 1 and working how how it changed as

Ex.3

a result. The commutator Eq. (3.42) is effectively the result of taking a vector from A to C via B and via D, and asking how the two resulting vectors are different.

The Riemann tensor has a number of symmetries. In a locally inertial frame,

$$R^{\alpha}{}_{\beta\mu\nu} = \tfrac{1}{2} g^{\alpha\sigma} (g_{\sigma\nu,\beta\mu} - g_{\sigma\mu,\beta\nu} + g_{\beta\mu,\sigma\nu} - g_{\beta\nu,\sigma\mu}), \tag{3.44}$$

and so

$$R_{\alpha\beta\mu\nu} \equiv g_{\alpha\sigma} R^{\sigma}{}_{\beta\mu\nu} = \tfrac{1}{2} (g_{\alpha\nu,\beta\mu} - g_{\alpha\mu,\beta\nu} + g_{\beta\mu,\alpha\nu} - g_{\beta\nu,\alpha\mu}). \tag{3.45}$$

Note that this is *not* a tensor equation, even in these coordinates: in such inertial coordinates $V^{\alpha}{}_{,\beta} = V^{\alpha}{}_{;\beta}$ and so an expression involving single partial derivatives of inertial coordinates can be trivially rewritten as a (covariant) tensor equation by simply rewriting the commas as semicolons; however the same is *not* true of second derivatives, so that Eq. (3.45) does not trivially correspond to a covariant expression.

By simply permuting indexes in Eq. (3.45), you can see that

$$R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta} \tag{3.46a}$$

$$R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\alpha\mu\nu\beta} = 0. \tag{3.46b}$$

These are tensor equations so that (as usual) although we worked them out in a particular coordinate system, they are true in all coordinate systems, and tell us about the symmetry properties of the underlying geometrical object.

5.2 Geodesic deviation

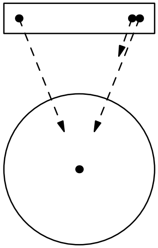

Figure 2

In Sect. 2.1 of part 1, we briefly imagined two objects in free fall near the earth (Fig. 2), and noted that the distance between them would decrease as they both moved towards the centre of the earth. We are now able to state that these free-falling objects are following geodesics in the spacetime surrounding the earth, which is curved as a result of the earth's mass (though we cannot say much more than this without doing the calculation, which is a bit of physics we do not know before the next part). We see, then, that the effect of the curvature of spacetime is to cause the distance between these two geodesics to decrease; this is known as *geodesic deviation*, and we are now in a position to see how it relates to curvature.

geod
tion

Schutz covers this at the end of his section 6.5. I plan to describe it a different way, partly because I find his explanation somewhat confusing, but also because a more geometrically-minded explanantion makes rather a change from continuous components.

First, some useful formulae. Marginally rewriting Eq. (3.43), we find

$$[\nabla_{\bar{X}}, \nabla_{\bar{Y}}]\bar{V} = X^\beta Y^\nu [\nabla_\beta, \nabla_\nu] V^\mu \bar{e}_\mu = R^\mu{}_{\alpha\beta\nu} V^\alpha X^\beta Y^\nu \bar{e}_\mu. \tag{3.47}$$

Next, using the commutator $[\bar{A}, \bar{B}] \equiv \bar{A}\bar{B} - \bar{B}\bar{A}$, we find

$$\nabla_{\bar{A}}\bar{B} - \nabla_{\bar{B}}\bar{A} = [\bar{A}, \bar{B}], \tag{3.48}$$

which is proved in example 3.15.

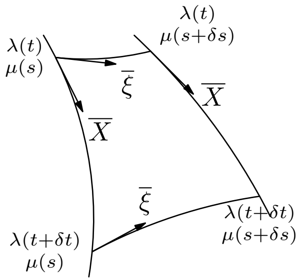

Figure 3

Consider two sets of curves, $\lambda(t)$ corresponding to a field of tangent vectors $\bar{X}$, and $\mu(s)$ with tangent vectors $\bar{\xi}$, and suppose that, in some region of the manifold, they cross each other (see Fig. 3). Choose the curves and their parameterisation such that each of the λ curves is a curve of constant s and each of the μ curves is a curve of constant t. Thus, specifically, the $\bar{\xi}$ vector – known as the 'connecting vector' – joins points on two λ curves which have the same parameter t. What we have actually described, here, is (part of) a set of coordinate functions; you will see that the curves λ and μ have exactly the properties that the conventionally-written coordinate functions x^α have. Because of this construction, it does not matter which order we take the derivatives $\mathrm{d}/\mathrm{d}t$ and $\mathrm{d}/\mathrm{d}s$, so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t} \Leftrightarrow \left[\frac{\mathrm{d}}{\mathrm{d}t}, \frac{\mathrm{d}}{\mathrm{d}s}\right] = 0,$$

or, since $\bar{X} = \mathrm{d}/\mathrm{d}t$ and $\bar{\xi} = \mathrm{d}/\mathrm{d}s$,

$$[\bar{X}, \bar{\xi}] = 0.$$

Thus, referring to Eq. (3.48),

$$\nabla_{\bar{X}}\bar{\xi} = \nabla_{\bar{\xi}}\bar{X}. \tag{3.49}$$

Now suppose particularly that the curves $\lambda(t)$ are geodesics, which means that $\nabla_{\bar{X}}\bar{X} = 0$. Then the vector $\bar{\xi}$ joins points on the two geodesics which have the same affine parameter. That means that the second derivative of $\bar{\xi}$ carries information about how quickly the two geodesics are accelerating apart (note that this is 'acceleration' in the sense of 'second derivative of position coordinate', and not anything that would be measured by an accelerometer – observers on the two geodesics would of course experience *zero* physical acceleration). With the work above, the calculation is easy. The second derivative is

$$\nabla_{\bar{X}}\nabla_{\bar{X}}\bar{\xi} = \nabla_{\bar{X}}\nabla_{\bar{\xi}}\bar{X} = \nabla_{\bar{\xi}}\nabla_{\bar{X}}\bar{X} + R^{\mu}{}_{\alpha\beta\nu}X^{\alpha}X^{\beta}\xi^{\nu}\bar{e}_{\mu}, \tag{3.50}$$

where the first equality comes from Eq. (3.49) and the second from Eq. (3.47). The first term on the right-hand side disappears since $\nabla_{\bar{X}}\bar{X} = 0$ along a geodesic. Now, the covariant derivative with respect to the vector $\bar{X}$ is just the derivative with respect to the geodesic's parameter t (since λ is part of a coordinate system, see Sect. 2.2), so that this equation turns into

$$\left(\frac{\mathrm{d}^2\bar{\xi}}{\mathrm{d}t^2}\right)^{\mu} = R^{\mu}{}_{\alpha\beta\nu}X^{\alpha}X^{\beta}\xi^{\nu}. \tag{3.51}$$

Thus the amount by which two geodesics diverge depends on the curvature of the space they are passing through. Note that the left-hand side here is the μ-component of the second derivative of the vector $\bar{\xi}$, and is a conventional shortcut for $\nabla_{\bar{X}}\nabla_{\bar{X}}$; it is *not* the second derivative of the ξ^{μ} component $\mathrm{d}^2\xi^{\mu}/\mathrm{d}t^2$, though some books (eg, [1, §1.9]) rather confusingly write it this way. Ex.3
Ex.3

References

[1] John Stewart. *Advanced General Relativity*. Cambridge, 1991.

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

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

Examples

Some of the examples below are taken from the earlier presentations of this course by Martin Hendry; these are noted by “[MAH]” and the example sheet where they occurred. Most of the exercises in Schutz’s §6.8 should be accessible.

Example 3.1 (section 1.3)

By considering the contraction of the gradient with a vector $a\,\mathrm{d}/\mathrm{d}t + b\,\mathrm{d}/\mathrm{d}s$, show that the gradient one-form defined by Eq. (3.5) is a linear function of its argument, and therefore a valid one-form.

Example 3.2 (section 1.4)

In the $\{x, y\}$ cartesian coordinate system, the metric is simply $\mathrm{diag}(1, 1)$. Consider a new coordinate system $\{u, v\}$, defined by

$$\begin{aligned} u &= \tfrac{1}{2}(x^2 - y^2) \\ v &= xy. \end{aligned} \tag{i}$$

(a) Write $x^1 = x$, $x^2 = y$, $x^{\bar{1}} = u$, $x^{\bar{2}} = v$, and thus calculate the numbers $\Lambda^{\bar{\imath}}{}_{j}$ and $\Lambda^{i}{}_{\bar{\jmath}}$.

(b) From Eq. (2.16),

$$g_{\bar{\imath}\bar{\jmath}} = \Lambda^{i}{}_{\bar{\imath}} \Lambda^{j}{}_{\bar{\jmath}} g_{ij}.$$

Thus calculate the components $g_{\bar{\imath}\bar{\jmath}}$ of the metric in terms of the coordinates $\{u, v\}$.

(c) A one-form has cartesian coordinates (A_x, A_y) and coordinates (A_u, A_v) in the new coordinate system. Show that

$$A_u = \frac{xA_x - yA_y}{x^2 + y^2},$$

and derive the corresponding expression for A_v. [MAH, 3.7; Objective 1].

Example 3.3 (section 1.4)

(a) Write down the expressions for cartesian coordinates $\{x, y\}$ as functions of polar coordinates $\{r, \theta\}$, thus calculate $\partial x/\partial r$, $\partial x/\partial\theta$, $\partial y/\partial r$ and $\partial y/\partial\theta$, and thus find the components of the transformation matrix from cartesian to polar coordinates, Eq. (3.8b).

(b) The inverse transformation is

$$r^2 = x^2 + y^2, \qquad \theta = \arctan\left(\frac{y}{x}\right).$$

Differentiate these, and thus obtain the inverse transformation matrix Eq. (3.8a). Verify that the product of these two matrices is indeed the identity matrix. Compare Sect. 3.2 of part 2.

(c) Let $\bar{V}$ be a vector with cartesian coordinates $\{x, y\}$, so that

$$\bar{V} = x\bar{e}_x + y\bar{e}_y.$$

Show that $\dot{\bar{V}}$ and $\ddot{\bar{V}}$ have components $\{\dot{x}, \dot{y}\}$ and $\{\ddot{x}, \ddot{y}\}$ in this basis.

(d) Using the relations $x = r\cos\theta$ and $y = r\sin\theta$, write down expressions for $\dot{x}$, $\dot{y}$, $\ddot{x}$ and $\ddot{y}$ in terms of polar coordinates r and θ and their time derivatives.

(e) Now use the general transformation law Eq. (3.8a)

$$V^{\bar{\imath}} = \Lambda^{\bar{\imath}}_{\ j} V^j = \frac{\partial x^{\bar{\imath}}}{\partial x^j} V^j$$

to transform the components of the vectors $\dot{\bar{V}}$ and $\ddot{\bar{V}}$ which you obtained in (c), into the polar basis $\{\bar{e}_r, \bar{e}_\theta\}$, and show that

$$\dot{\bar{V}} = \dot{r}\bar{e}_r + \dot{\theta}\bar{e}_\theta$$

$$\ddot{\bar{V}} = \left(\ddot{r} - r\dot{\theta}^2\right)\bar{e}_r + \left(\ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta}\right)\bar{e}_\theta.$$

[MAH, 2.1; Objective 1]

Example 3.4 (section 1.4)

Define a scalar field, ϕ, by

$$\phi(x, y) = x^2 + y^2 + 2xy,$$

for cartesian coordinates $\{x, y\}$.

(a) From Eq. (3.5), the i-th component of the gradient one-form $\widetilde{\mathrm{d}}\phi$ is obtained by taking the contraction of the gradient with the basis vector $\bar{e}_i = \partial/\partial x^i$. Thus write down the components of the gradient one-form with respect to the cartesian basis.

(b) The result of example 2.7 of part 2 says that the transformation law for the components of a one-form is

$$A_{\bar{\imath}} = \Lambda^j_{\bar{\imath}} A_j = \frac{\partial x^j}{\partial x^{\bar{\imath}}} A_j.$$

Thus determine the components of $\widetilde{\mathrm{d}}\phi$ in polar coordinates $\{r, \theta\}$.

(c) By expressing ϕ in terms of r and θ, obtain directly the polar components of $\widetilde{\mathrm{d}}\phi$ and verify that they agree with those obtained in (b).

(d) Write down the components of the metric tensor in cartesian coordinates, g_{xx}, g_{xy}, g_{yx}, g_{yy}, and by examining Eq. (2.9), write down the components of the metric tensor with raised indexes, g^{xx}, g^{xy}, g^{yx}, g^{yy}. Hence determine the cartesian components of the *vector* gradient $\overline{\mathrm{d}}\phi$ (ie, with raised index).

(e) Recall the metric for polar coordinates, and thus the components g^{rr}, $g^{r\theta}$, $g^{\theta r}$ and $g^{\theta\theta}$. Hence determine the polar components of $\overline{\mathrm{d}}\phi$. Comment on the answers to parts (d) and (e). [MAH, 2.2]

Example 3.5 (section 2.2)

Consider a vector field $\bar{V}$ with cartesian components $\{V^x, V^y\} = \{x^2 + 3y, y^2 + 3x\}$.

(a) Using the transformation law for a $\binom{1}{0}$ tensor, and the result of example 3.3, determine $\{V^r, V^\theta\}$, the components of the same vector field $\bar{V}$ with respect to the polar basis $\{\bar{e}_r, \bar{e}_\theta\}$.

(b) Write down the components of the covariant derivative $V^i{}_{;j}$ in cartesian coordinates.

(c) Using the fact that $V^i{}_{;j}$ transforms as a $\binom{1}{1}$ tensor, compute the components of the covariant derivative with respect to the polar coordinate basis by transforming the $V^i{}_{;j}$ obtained in part (b).

(d) Now, taking a different tack, compute the polar components of the covariant derivative of $\bar{V}$, by differentiating the polar coordinates obtained in (a). That is, use Eq. (3.15b) and the Christoffel symbols for polar coordinates, Eq. (3.13).

(e) Verify that the polar components obtained in (c) and (d) are the same. [MAH, 4.5]

Example 3.6 (section 2.2)

Do example 3.5 again, but this time working with the one-form field $\widetilde{A}$, with cartesian components $\{x^2 + 3y, y^2 + 3x\}$. [MAH, 4.6]

Example 3.7 (section 2.2)

Comparing example 3.5 and example 3.6, verify that in both cartesian and polar coordinates

$$g_{ik} V^k{}_{;j} = A_{i;j}.$$

[MAH, 4.7]

Example 3.8 (section 2.3)

Let A_j be the components of an arbitrary one-form. Write down the transformation law for A_j and for its covariant derivative $A_{j;k}$. By considering the expression for $A_{\bar{j};\bar{k}}$, in a transformed coordinate system, show that the transformation law for the Christoffel symbols has the form

$$\Gamma^{\bar{i}}_{\bar{j}\bar{k}} = \frac{\partial x^{\bar{i}}}{\partial x^i} \frac{\partial x^j}{\partial x^{\bar{j}}} \frac{\partial x^k}{\partial x^{\bar{k}}} \Gamma^i_{jk} + \frac{\partial x^{\bar{i}}}{\partial x^l} \frac{\partial^2 x^l}{\partial x^{\bar{j}} \partial x^{\bar{k}}}.$$

The fact that this does *not* look anything like Eq. (2.16) further demonstrates that the Christoffel symbols are not the components of a tensor. [MAH, 4.1]

Example 3.9 (section 2.3)

Suppose that in one coordinate system the Christoffel symbols are symmetric in their lower indexes, $\Gamma^i_{\ jk} = \Gamma^i_{kj}$. By considering the transformation law for the Christoffel symbols, obtained in example 3.8, show that they will be symmetric in any coordinate system. [MAH, 4.2]

Example 3.10 (section 3.2)

Things to think about: Why have you never had to learn about covariant differentiation before now? The glib answer is, of course, that you weren't learning GR; but what was it about the vector calculus that you did learn that meant you never had to know about connection coefficients? Or, given that you did effectively learn about them, but didn't know that was what they were called, why do we have to go into so much more detail about them now? There are a variety of answers to these questions, at different levels.

Example 3.11 (section 4)

(a) On the surface of a sphere, we can pick coordinates θ and ϕ, where θ is the colatitude, and ϕ is the azimuthal coordinate. The components of the metric in these coordinates are

$$g_{\theta\theta} = 1, \qquad g_{\phi\phi} = \sin^2\theta, \qquad \text{others zero.}$$

Show that the components of the metric with raised indexes are

$$g^{\theta\theta} = 1, \qquad g^{\phi\phi} = \frac{1}{\sin^2\theta}, \qquad \text{others zero.}$$

[4 marks]

(b) The Christoffel symbols are defined as

$$\Gamma^\alpha_{\mu\nu} = \tfrac{1}{2} g^{\alpha\beta}(g_{\beta\mu,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta}),$$

and the geodesic equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}t}\right) + \Gamma^\alpha_{\beta\mu}\frac{\mathrm{d}x^\beta}{\mathrm{d}t}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = 0,$$

for a geodesic with parameter t. Using these find the Christoffel symbols for these coordinates (ie, $\Gamma^\theta_{\theta\theta}$, $\Gamma^\theta_{\theta\phi}$ and so on), and thus show that the geodesic equations for these coordinates are

$$\ddot\theta - \sin\theta\cos\theta\,\dot\phi^2 = 0 \qquad \text{(i)}$$

$$\ddot\phi + 2\frac{\cos\theta}{\sin\theta}\dot\theta\dot\phi = 0, \qquad \text{(ii)}$$

where dots indicate differentiation with respect to the parameter t. [12 marks]

(c) Using the result of part (b), or any other properties of geodesics which you know, explain, giving reasons, which of the following curves are geodesics, for affine parameter t.

1. $\phi = t, \theta = \pi/2$ 2. $\phi = t, \theta = \pi/4$ 3. $\phi = t, \theta = 0$
4. $\phi = t, \theta = t$ 5. $\phi = \phi_0, \theta = t$ 6. $\phi = \phi_0, \theta = 2t - 1$
7. $\phi = \phi_0, \theta = t^2$

[6 marks]

(d) If $\bar{U}$ is the tangent vector to a geodesic, so that

$$\nabla_{\bar{U}}\bar{U} = 0,$$

prove that $\bar{V} = a\bar{U}$ is also tangent to a geodesic, for any constant number a. [8 marks]

[MAH 5.1, originally, later Class Test 2002, with additions; Objective 2, Objective 5]

Example 3.12 (section 5.1)

Prove Eq. (3.43). Write $\nabla_\alpha \nabla_\beta V^\mu = \nabla_\alpha (V^\mu{}_{;\beta}) = (V^\mu{}_{;\beta})_{;\alpha}$, and use the expression Eq. (3.21) to expand the derivative with resepect to x^α. At this point, decide to work in LIF coordinates, in which all the $\Gamma^\mu_{\alpha\beta} = 0$, making the algebra easier. Thus deduce that $\nabla_\alpha \nabla_\beta V^\mu = V^\mu{}_{,\beta\alpha} + \Gamma^\mu_{\nu\beta,\alpha} V^\nu$. You can then immediately write down an expression for $\nabla_\beta \nabla_\alpha V^\mu$. Subtract these two expressions (to form $[\nabla_\alpha, \nabla_\beta] V^\mu$), noting that the usual partial differentiation of components commutes: $V^\mu{}_{,\alpha\beta} = V^\mu{}_{,\beta\alpha}$. Compare the result with the definition of the Riemann tensor in Eq. (3.40), and arrive at Eq. (3.43). If you get stuck with the algebra, the proof's in Schutz §6.5.

Example 3.13 (section 5.1)

Prove Eq. (3.44). Recall that in a locally inertial frame, at a point P, the components of the connection are zero: $\Gamma^{\alpha}_{\mu\nu} = 0$. The derivatives, however, are not, so use Eq. (3.32) straightforwardly to find $\Gamma^{\alpha}_{\mu\nu,\sigma}$. Recall Eq. (3.28a), and that partial derivatives always commute.

Example 3.14 (section 5.1)

In example 3.11 you calculated the Christoffel symbols for the surface of the unit sphere. Calculate the components of the curvature tensor for these coordinates, plus the Ricci tensor $R_{\beta\gamma} = R^{\alpha}{}_{\beta\alpha\gamma}$ and the Ricci scalar $R = g^{\beta\gamma} R_{\beta\gamma}$ (see part 4).

You can most conveniently do this by calculating selected components of the curvature tensor $R_{\alpha\beta\mu\nu}$ obtained by lowering the first index on Eq. (3.40); you can cut down the number of calculations you need to do by using the symmetry relations Eq. (3.46) heavily. Why should you not use Eq. (3.45), which appears to be more straightforward?

This question is long-winded rather than terribly hard. It's worthwhile slogging through it, however, since it gives valuable practice handling indices, and makes the idea of the curvature tensor rather more tangible. [Objective 6]

Example 3.15 (section 5.2)

Prove Eq. (3.48), by writing it in component form. Recall Eq. (3.27). The last step is the tricky bit, but recall that for a (tangent) vector $\bar{A}$, $\bar{A} f = A^\mu \bar{e}_\mu f = A^\mu f_{,\mu}$, where f is any function, including a vector component.

Example 3.16 (section 5.2)

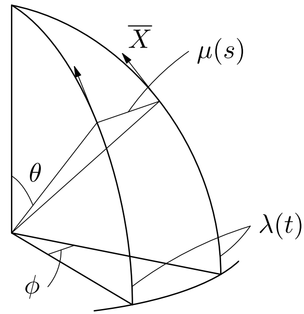

Figure 4

Consider coordinates on a sphere, as you did in example 3.11, and consider the geodesics $\lambda(t)$ in Fig. 4 with affine parameter t and tangent vectors $\bar{X}$ – these are great circles through the poles. The curves $\mu(s)$ with tangent vectors $\bar{\xi}$ are connecting curves as discussed in Sect. 5.2.

We can parameterise the curve $\lambda(t)$ using the coordinates (θ, ϕ), as

$$\lambda(t) : \theta(\lambda(t)) = t; \qquad \phi(\lambda(t)) = \phi_0$$

(compare Sect. 1.2), and you verified in example 3.11 that this does indeed satisfy the geodesic equation.

(a) Using Eq. (3.1), show that the components of $\bar{X}$ are

$$X^\theta = 1, \qquad X^\phi = 0.$$

(b) Write Eq. (3.51) as

$$g_{\lambda\mu}(\nabla_{\bar{X}}\nabla_{\bar{X}}\bar{\xi})^{\mu} - g_{\lambda\mu}R^{\mu}{}_{\alpha\beta\nu}X^{\alpha}X^{\beta}\xi^{\nu} = 0 \qquad \text{(i)}$$

and, by using the components of the curvature tensor which you worked out in example 3.14, show that

$$(\nabla_{\bar{X}}\nabla_{\bar{X}}\bar{\xi})^{\theta} = 0 \qquad \text{(iia)}$$

$$(\nabla_{\bar{X}}\nabla_{\bar{X}}\bar{\xi})^{\phi} + \xi^{\phi} = 0. \qquad \text{(iib)}$$

This tells us that the connecting vector – the tangent vector to the family of curves $\mu(s)$, connecting points of equal affine parameter along the geodesics $\lambda(t)$ – does not change its θ component, but does change its ϕ component. Which isn't much of a surprise.

(c) Can we get more out of this? Yes, but to do that we have to calculate $\nabla_{\bar{X}}\nabla_{\bar{X}}\bar{\xi}$, which isn't quite as challenging as it might look. From Eq. (3.18) we write

$$\nabla_{\bar{X}}\bar{\xi} = X^{\alpha}\nabla_{\alpha}\bar{\xi} = X^{\alpha}\bar{e}_{\beta}\xi^{\beta}{}_{;\alpha} = X^{\alpha}\bar{e}_{\beta}\left(\frac{\partial\xi^{\beta}}{\partial x^{\alpha}} + \Gamma^{\beta}_{\alpha\gamma}\xi^{\gamma}\right). \qquad \text{(iii)}$$

You have worked out the Christoffel symbols for these coordinates in example 3.11, so we *could* trundle on through this calculation, and find expressions for the components of the connecting vector $\bar{\xi}$ from Eq. (ii). In order to illustrate something useful in a reasonable amount of time, however, we will short-circuit that by using our previous knowledge of this coordinate system.

The curve

$$\mu(s) : \theta(s) = \theta_0, \quad \phi(s) = s$$

is *not* a geodesic (it is a small circle at colatitude θ_0), but it does connect points on the geodesics $\lambda(t)$ with equal affine parameter t; it is a connecting curve for this family of geodesics.

Convince yourself of this and, as in part (a) above, satisfy yourself that the tangent vector to this curve, $\xi = \mathrm{d}/\mathrm{d}s$, has components $\xi^\theta = 0$ and $\xi^\phi = 1$; and use this together with the components of the tangent vector $\bar{X}$ and the expression Eq. (iii) to deduce that

$$\dot{\bar{\xi}} \equiv \nabla_{\bar{X}} \bar{\xi} = 0 \bar{e}_\theta + \cot\theta \, \bar{e}_\phi,$$

(where $\dot{\bar{\xi}}$ is simply a convenient – and conventional – notation for $\nabla_{\bar{X}} \bar{\xi}$) or $\dot{\xi}^\theta = 0$, $\dot{\xi}^\phi = \cot\theta$.

(d) So far so good. In exactly the same way, take the covariant derivative of $\dot{\bar{\xi}}$, and discover that

$$\nabla_{\bar{X}} \dot{\bar{\xi}} = \nabla_{\bar{X}} \nabla_{\bar{X}} \bar{\xi} = 0 \bar{e}_\theta - 1 \bar{e}_\phi = -\bar{\xi},$$

and note that this $\bar{\xi}$ does in fact accord with the geodesic deviation equation of Eq. (ii).

Note that this example is somewhat fake, in that, in (c), we set up the curve $\mu(s)$ as a connecting curve, and all we have done here is verify that this was consistent. If we were doing this for real, we would not know (all of) the components of $\bar{\xi}$ beforehand, but would carry on differentiating $\bar{\xi}$ as we started to do in (c), put the result into the differential equation Eq. (ii) and thus deduce expressions for the components ξ^μ.

As a final point, note that the *length* of the connecting vector $\bar{\xi}$ is just

$$\mathsf{g}(\bar{\xi}, \bar{\xi}) = g_{\alpha\beta} \xi^\alpha \xi^\beta = \sin^2\theta,$$

which you could possibly have worked out from school trignometry (but it wouldn't have been half so much fun).

[Objective 5, Objective 7]

Example 3.17 (section 5.2)

(a) Describe the relationship between geodesic deviation and the metric of a spacetime, referring in your answer both to the equation of geodesic deviation (Eq. (ii) below), and to the behaviour of test particles in free fall near the Earth's surface. [5]

(b) In the newtonian limit, the metric can be written as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{i}$$

where

$$\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$$
$$h_{\mu\nu} = \begin{cases} -2\phi & \mu = \nu \\ 0 & \mu \neq \nu \end{cases},$$

and ϕ is the newtonian gravitational potential $\phi(r) = GM/r$. In this limit, and with this metric, the curvature tensor can be written as

$$2R_{\alpha\beta\mu\nu} = h_{\alpha\nu,\beta\mu} + h_{\beta\mu,\alpha\nu} - h_{\alpha\mu,\beta\nu} - h_{\beta\nu,\alpha\mu}.$$

The equation for geodesic deviation is

$$\frac{\mathrm{d}^2\xi^\alpha}{\mathrm{d}t^2} = R^\alpha{}_{\beta\mu\nu} U^\beta U^\mu \xi^\nu, \tag{ii}$$

where the vectors $\bar{U}$ are tangent to geodesics, and we can take them to be velocity vectors.

Consider two particles in free fall just above the Earth's north pole, so that their (cartesian) coordinates are both approximately $x = y = 0$, $z = R$, where R is the radius of the Earth. Take them to be separated by a separation vector $\bar{\xi} = (0, \xi^x, 0, 0)$, where $\xi^x \ll R$. Since they are falling along geodesics, their velocity vectors are both approximately $\bar{U} = (U^t, 0, 0, U^z)$.

With this information, show that the two particles accelerate towards each other such that

$$\frac{d^2\xi^x}{dt^2} = -\frac{GM}{r^3}\xi^x \tag{iii}$$

to first order in ϕ (given values for G, M and R, why can we take $\phi^2 \ll 0$?).

[Since these are non-relativistic particles, you may assume, at the appropriate point, that $|U^t| \gg |U^z|$, and thus that $|U^t|^2 \approx -1$.] [20]

(c) If we had used a different metric to describe the same newtonian spacetime, rather than that in Eq. (i), would we have obtained a different result for the geodesic deviation, Eq. (iii)? Explain your answer. [5] [Objective 5, Objective 7; based on degree exam, 2003]