Massachusetts Institute of Technology
Department of Physics

Physics 8.962 Spring 1999

Introduction to Tensor Calculus for General Relativity

1 Introduction

There are three essential ideas underlying general relativity (GR). The first is that spacetime may be described as a curved, four-dimensional mathematical structure called a pseudo-Riemannian manifold. In brief, time and space together comprise a curved four-dimensional non-Euclidean geometry. Consequently, the practitioner of GR must be familiar with the fundamental geometrical properties of curved spacetime. In particular, the laws of physics must be expressed in a form that is valid independently of any coordinate system used to label points in spacetime.

The second essential idea underlying GR is that at every spacetime point there exist locally inertial reference frames, corresponding to locally flat coordinates carried by freely falling observers, in which the physics of GR is locally indistinguishable from that of special relativity. This is Einstein's famous strong equivalence principle and it makes general relativity an extension of special relativity to a curved spacetime. The third key idea is that mass (as well as mass and momentum flux) curves spacetime in a manner described by the tensor field equations of Einstein.

These three ideas are exemplified by contrasting GR with Newtonian gravity. In the Newtonian view, gravity is a force accelerating particles through Euclidean space, while time is absolute. From the viewpoint of GR, there is no gravitational force. Rather, in the absence of electromagnetic and other forces, particles follow the straightest possible paths (geodesics) through a spacetime curved by mass. Freely falling particles define locally inertial reference frames. Time and space are not absolute but are combined into the four-dimensional manifold called spacetime.

Working with GR, particularly with the Einstein field equations, requires some understanding of differential geometry. In these notes we will develop the essential mathematics needed to describe physics in curved spacetime. Many physicists receive their introduction to this mathematics in the excellent book of Weinberg (1972). Weinberg minimizes the geometrical content of the equations by representing tensors using component notation. We believe that it is equally easy to work with a more geometrical description, with the additional benefit that geometrical notation makes it easier to distinguish physical results that are true in any coordinate system (e.g., those expressible using vectors) from those that are dependent on the coordinates. Because the geometry of spacetime is so intimately related to physics, we believe that it is better to highlight the geometry from the outset. In fact, using a geometrical approach allows us to develop the essential differential geometry as an extension of vector calculus. Our treatment is closer to that Wald (1984) and closer still to Misner, Thorne and Wheeler (1973). These books are rather advanced. For the newcomer to general relativity we warmly recommend Schutz (1985). Our notation and presentation is patterned largely after Schutz. The student wishing additional practice problems in GR should consult Lightman *et al.* (1975). A slightly more advanced mathematical treatment is provided in the excellent notes of Carroll (1997).

These notes assume familiarity with special relativity. We will adopt units in which the speed of light $c = 1$. Greek indices (μ, ν, etc., which take the range $\{0, 1, 2, 3\}$) will be used to represent components of tensors. The Einstein summation convention is assumed: repeated upper and lower indices are to be summed over their ranges, e.g., $A^\mu B_\mu \equiv A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3$. Four-vectors will be represented with an arrow over the symbol, e.g., $\vec{A}$, while one-forms will be represented using a tilde, e.g., $\tilde{B}$. Spacetime points will be denoted in boldface type; e.g., $\mathbf{x}$ refers to a point with coordinates x^μ. Our metric has signature +2; the flat spacetime Minkowski metric components are $\eta_{\mu\nu} = \mathbf{diag}(-1, +1, +1, +1)$.

2 Vectors and one-forms

The essential mathematics of general relativity is differential geometry, the branch of mathematics dealing with smoothly curved surfaces (differentiable manifolds). The physicist does not need to master all of the subtleties of differential geometry in order to use general relativity. (For those readers who want a deeper exposure to differential geometry, see the introductory texts of Lovelock and Rund 1975, Bishop and Goldberg 1980, or Schutz 1980.) It is sufficient to develop the needed differential geometry as a straightforward extension of linear algebra and vector calculus. However, it is important to keep in mind the geometrical interpretation of physical quantities. For this reason, we will not shy from using abstract concepts like points, curves and vectors, and we will distinguish between a vector $\vec{A}$ and its components A^μ. Unlike some other authors (e.g., Weinberg 1972), we will introduce geometrical objects in a coordinate-free manner, only later introducing coordinates for the purpose of simplifying calculations. This approach requires that we distinguish vectors from the related objects called one-forms. Once the differences and similarities between vectors, one-forms and tensors are clear, we will adopt a unified notation that makes computations easy.

2.1 Vectors

We begin with vectors. A vector is a quantity with a magnitude and a direction. This primitive concept, familiar from undergraduate physics and mathematics, applies equally in general relativity. An example of a vector is $d\vec{x}$, the difference vector between two infinitesimally close points of spacetime. Vectors form a linear algebra (i.e., a vector space). If $\vec{A}$ is a vector and a is a real number (scalar) then $a\vec{A}$ is a vector with the same direction (or the opposite direction, if $a < 0$) whose length is multiplied by $|a|$. If $\vec{A}$ and $\vec{B}$ are vectors then so is $\vec{A} + \vec{B}$. These results are as valid for vectors in a curved four-dimensional spacetime as they are for vectors in three-dimensional Euclidean space.

Note that we have introduced vectors without mentioning coordinates or coordinate transformations. Scalars and vectors are invariant under coordinate transformations; vector components are not. The whole point of writing the laws of physics (e.g., $\vec{F} = m\vec{a}$) using scalars and vectors is that these laws do not depend on the coordinate system imposed by the physicist.

We denote a spacetime point using a boldface symbol: $\mathbf{x}$. (This notation is *not* meant to imply coordinates.) Note that $\mathbf{x}$ refers to a point, not a vector. In a curved spacetime the concept of a radius vector $\vec{x}$ pointing from some origin to each point $\mathbf{x}$ is not useful because vectors defined at two different points cannot be added straightforwardly as they can in Euclidean space. For example, consider a sphere embedded in ordinary three-dimensional Euclidean space (i.e., a two-sphere). A vector pointing east at one point on the equator is seen to point radially outward at another point on the equator whose longitude is greater by 90°. The radially outward direction is undefined on the sphere.

Technically, we are discussing *tangent vectors* that lie in the *tangent space* of the manifold at each point. For example, a sphere may be embedded in a three-dimensional Euclidean space into which may be placed a plane tangent to the sphere at a point. A two-dimensional vector space exists at the point of tangency. However, such an embedding is not required to define the tangent space of a manifold (Walk 1984). As long as the space is smooth (as assumed in the formal definition of a manifold), the difference vector $d\vec{x}$ between to infinitesimally close points may be defined. The set of all $d\vec{x}$ defines the tangent space at $\mathbf{x}$. By assigning a tangent vector to every spacetime point, we can recover the usual concept of a vector field. However, without additional preparation one cannot compare vectors at different spacetime points, because they lie in different tangent spaces. In Section 5 we introduce parallel transport as a means of making this comparison. Until then, we consider only tangent vectors at $\mathbf{x}$. To emphasize the status of a tangent vector, we will occasionally use a subscript notation: $\vec{A}_X$.

2.2 One-forms and dual vector space

Next we introduce one-forms. A one-form is defined as a linear scalar function of a vector. That is, a one-form takes a vector as input and outputs a scalar. For the one-form $\tilde{P}$, $\tilde{P}(\vec{V})$ is also called the scalar product and may be denoted using angle brackets:

$$\tilde{P}(\vec{V}) = \langle \tilde{P}, \vec{V} \rangle \ . \tag{1}$$

The one-form is a linear function, meaning that for all scalars a and b and vectors $\vec{V}$ and $\vec{W}$, the one-form $\tilde{P}$ satisfies the following relations:

$$\tilde{P}(a\vec{V} + b\vec{W}) = \langle \tilde{P}, a\vec{V} + b\vec{W} \rangle = a\langle \tilde{P}, \vec{V} \rangle + b\langle \tilde{P}, \vec{W} \rangle = a\tilde{P}(\vec{V}) + b\tilde{P}(\vec{W}) \ . \tag{2}$$

Just as we may consider any function $f(\ )$ as a mathematical entity independently of any particular argument, we may consider the one-form $\tilde{P}$ independently of any particular vector $\vec{V}$. We may also associate a one-form with each spacetime point, resulting in a one-form field $\tilde{P} = \tilde{P}_{\mathbf{X}}$. Now the distinction between a point a vector is crucial: $\tilde{P}_{\mathbf{X}}$ is a one-form at point $\mathbf{x}$ while $\tilde{P}(\vec{V})$ is a scalar, defined implicitly at point $\mathbf{x}$. The scalar product notation with subscripts makes this more clear: $\langle \tilde{P}_{\mathbf{X}}, \vec{V}_{\mathbf{X}} \rangle$.

One-forms obey their own linear algebra distinct from that of vectors. Given any two scalars a and b and one-forms $\tilde{P}$ and $\tilde{Q}$, we may define the one-form $a\tilde{P} + b\tilde{Q}$ by

$$(a\tilde{P} + b\tilde{Q})(\vec{V}) = \langle a\tilde{P} + b\tilde{Q}, \vec{V} \rangle = a\langle \tilde{P}, \vec{V} \rangle + b\langle \tilde{Q}, \vec{V} \rangle = a\tilde{P}(\vec{V}) + b\tilde{Q}(\vec{V}) \ . \tag{3}$$

Comparing equations (2) and (3), we see that vectors and one-forms are linear operators on each other, producing scalars. It is often helpful to consider a vector as being a linear scalar function of a one-form. Thus, we may write $\langle \tilde{P}, \vec{V} \rangle = \tilde{P}(\vec{V}) = \vec{V}(\tilde{P})$. The set of all one-forms is a vector space distinct from, but complementary to, the linear vector space of vectors. The vector space of one-forms is called the *dual* vector (or cotangent) space to distinguish it from the linear space of vectors (tangent space).

Although one-forms may appear to be highly abstract, the concept of dual vector spaces is familiar to any student of quantum mechanics who has seen the Dirac bra-ket notation. Recall that the fundamental object in quantum mechanics is the state vector, represented by a ket $|\psi\rangle$ in a linear vector space (Hilbert space). A distinct Hilbert space is given by the set of bra vectors $\langle\phi|$. Bra vectors and ket vectors are linear scalar functions of each other. The scalar product $\langle\phi|\psi\rangle$ maps a bra vector and a ket vector to a scalar called a probability amplitude. The distinction between bras and kets is necessary because probability amplitudes are complex numbers. As we will see, the distinction between vectors and one-forms is necessary because spacetime is curved.

3 Tensors

Having defined vectors and one-forms we can now define tensors. A tensor of rank (m, n), also called a (m, n) tensor, is defined to be a scalar function of m one-forms and n vectors that is linear in all of its arguments. It follows at once that scalars are tensors of rank $(0, 0)$, vectors are tensors of rank $(1, 0)$ and one-forms are tensors of rank $(0, 1)$. We may denote a tensor of rank $(2, 0)$ by $\mathsf{T}(\tilde{P}, \tilde{Q})$; one of rank $(2, 1)$ by $\mathsf{T}(\tilde{P}, \tilde{Q}, \vec{A})$, etc. Our notation will not distinguish a $(2, 0)$ tensor T from a $(2, 1)$ tensor T, although a notational distinction could be made by placing m arrows and n tildes over the symbol, or by appropriate use of dummy indices (Wald 1984).

The scalar product is a tensor of rank $(1, 1)$, which we will denote I and call the identity tensor:

$$\mathsf{I}(\tilde{P}, \vec{V}) \equiv \langle \tilde{P}, \vec{V} \rangle = \tilde{P}(\vec{V}) = \vec{V}(\tilde{P}) \ . \tag{4}$$

We call I the identity because, when applied to a fixed one-form $\tilde{P}$ and *any* vector $\vec{V}$, it returns $\tilde{P}(\vec{V})$. Although the identity tensor was defined as a mapping from a one-form and a vector to a scalar, we see that it may equally be interpreted as a mapping from a one-form to the same one-form: $\mathsf{I}(\tilde{P}, \cdot) = \tilde{P}$, where the dot indicates that an argument (a vector) is needed to give a scalar. A similar argument shows that I may be considered the identity operator on the space of vectors $\vec{V}$: $\mathsf{I}(\cdot, \vec{V}) = \vec{V}$.

A tensor of rank (m, n) is linear in all its arguments. For example, for $(m = 2, n = 0)$ we have a straightforward extension of equation (2):

$$\mathsf{T}(a\tilde{P} + b\tilde{Q}, c\tilde{R} + d\tilde{S}) = ac\,\mathsf{T}(\tilde{P}, \tilde{R}) + ad\,\mathsf{T}(\tilde{P}, \tilde{S}) + bc\,\mathsf{T}(\tilde{Q}, \tilde{R}) + bd\,\mathsf{T}(\tilde{q}, \tilde{S}) \ . \tag{5}$$

Tensors of a given rank form a linear algebra, meaning that a linear combination of tensors of rank (m, n) is also a tensor of rank (m, n), defined by straightforward extension of equation (3). Two tensors (of the same rank) are equal if and only if they return the same scalar when applied to all possible input vectors and one-forms. Tensors of different rank cannot be added or compared, so it is important to keep track of the rank of each tensor. Just as in the case of scalars, vectors and one-forms, tensor fields $\mathsf{T}_{\mathbf{X}}$ are defined by associating a tensor with each spacetime point.

There are three ways to change the rank of a tensor. The first, called the tensor (or outer) product, combines two tensors of ranks (m_1, n_1) and (m_2, n_2) to form a tensor of rank $(m_1 + m_2, n_1 + n_2)$ by simply combining the argument lists of the two tensors and thereby expanding the dimensionality of the tensor space. For example, the tensor product of two vectors $\vec{A}$ and $\vec{B}$ gives a rank $(2, 0)$ tensor

$$\mathsf{T} = \vec{A} \otimes \vec{B} \ , \quad \mathsf{T}(\tilde{P}, \tilde{Q}) \equiv \vec{A}(\tilde{P})\,\vec{B}(\tilde{Q}) \ . \tag{6}$$

We use the symbol $\otimes$ to denote the tensor product; later we will drop this symbol for notational convenience when it is clear from the context that a tensor product is implied.

Note that the tensor product is non-commutative: $\vec{A} \otimes \vec{B} \neq \vec{B} \otimes \vec{A}$ (unless $\vec{B} = c\vec{A}$ for some scalar c) because $\vec{A}(\tilde{P})\,\vec{B}(\tilde{Q}) \neq \vec{A}(\tilde{Q})\,\vec{B}(\tilde{P})$ for all $\tilde{P}$ and $\tilde{Q}$. We use the symbol $\otimes$ to denote the tensor product of any two tensors, e.g., $\tilde{P} \otimes \mathsf{T} = \tilde{P} \otimes \vec{A} \otimes \vec{B}$ is a tensor of rank $(2, 1)$. The second way to change the rank of a tensor is by contraction, which reduces the rank of a (m, n) tensor to $(m-1, n-1)$. The third way is the gradient. We will discuss contraction and gradients later.

3.1 Metric tensor

The scalar product (eq. 1) requires a vector and a one-form. Is it possible to obtain a scalar from two vectors or two one-forms? From the definition of tensors, the answer is clearly yes. Any tensor of rank $(0, 2)$ will give a scalar from two vectors and any tensor of rank $(2, 0)$ combines two one-forms to give a scalar. However, there is a particular $(0, 2)$ tensor field $\mathsf{g}_{\mathbf{X}}$ called the metric tensor and a related $(2, 0)$ tensor field $\mathsf{g}_{\mathbf{X}}^{-1}$ called the inverse metric tensor for which special distinction is reserved. The metric tensor is a symmetric bilinear scalar function of two vectors. That is, given vectors $\vec{V}$ and $\vec{W}$, g returns a scalar, called the dot product:

$$\mathsf{g}(\vec{V}, \vec{W}) = \vec{V} \cdot \vec{W} = \vec{W} \cdot \vec{V} = \mathsf{g}(\vec{W}, \vec{V}) \ . \tag{7}$$

Similarly, g^{-1} returns a scalar from two one-forms $\tilde{P}$ and $\tilde{Q}$, which we also call the dot product:

$$\mathsf{g}^{-1}(\tilde{P}, \tilde{Q}) = \tilde{P} \cdot \tilde{Q} = \tilde{P} \cdot \tilde{Q} = \mathsf{g}^{-1}(\tilde{P}, \tilde{Q}) \ . \tag{8}$$

Although a dot is used in both cases, it should be clear from the context whether g or g^{-1} is implied. We reserve the dot product notation for the metric and inverse metric tensors just as we reserve the angle brackets scalar product notation for the identity tensor (eq. 4). Later (in eq. 40) we will see what distinguishes g from other $(0, 2)$ tensors and g^{-1} from other symmetric $(2, 0)$ tensors.

One of the most important properties of the metric is that it allows us to convert vectors to one-forms. If we forget to include $\vec{W}$ in equation (7) we get a quantity, denoted $\tilde{V}$, that behaves like a one-form:

$$\tilde{V}(\,\cdot\,) \equiv \mathsf{g}(\vec{V}, \,\cdot\,) = \mathsf{g}(\,\cdot\,, \vec{V}) \ , \tag{9}$$

where we have inserted a dot to remind ourselves that a vector must be inserted to give a scalar. (Recall that a one-form is a scalar function of a vector!) We use the same letter to indicate the relation of $\vec{V}$ and $\tilde{V}$.

Thus, the metric g is a mapping from the space of vectors to the space of one-forms: $\mathsf{g} : \vec{V} \to \tilde{V}$. By definition, the inverse metric g^{-1} is the inverse mapping: $\mathsf{g}^{-1} : \tilde{V} \to \vec{V}$.

(The inverse always exists for nonsingular spacetimes.) Thus, if $\tilde{V}$ is defined for any $\vec{V}$ by equation (9), the inverse metric tensor is defined by

$$\vec{V}(\,\cdot\,) \equiv \mathsf{g}^{-1}(\tilde{V},\,\cdot\,) = \mathsf{g}^{-1}(\,\cdot\,,\tilde{V})\ . \tag{10}$$

Equations (4) and (7)–(10) give us several equivalent ways to obtain scalars from vectors $\vec{V}$ and $\vec{W}$ and their associated one-forms $\tilde{V}$ and $\tilde{W}$:

$$\langle\tilde{V},\vec{W}\rangle = \langle\tilde{W},\vec{V}\rangle = \vec{V}\cdot\vec{W} = \tilde{V}\cdot\tilde{W} = \mathsf{I}(\tilde{V},\vec{W}) = \mathsf{I}(\tilde{W},\vec{V}) = \mathsf{g}(\vec{V},\vec{W}) = \mathsf{g}^{-1}(\tilde{V},\tilde{W})\,. \tag{11}$$

3.2 Basis vectors and one-forms

It is possible to formulate the mathematics of general relativity entirely using the abstract formalism of vectors, forms and tensors. However, while the geometrical (coordinate-free) interpretation of quantities should always be kept in mind, the abstract approach often is not the most practical way to perform calculations. To simplify calculations it is helpful to introduce a set of linearly independent basis vector and one-form fields spanning our vector and dual vector spaces. In the same way, practical calculations in quantum mechanics often start by expanding the ket vector in a set of basis kets, e.g., energy eigenstates. By definition, the dimensionality of spacetime (four) equals the number of linearly independent basis vectors and one-forms.

We denote our set of basis vector fields by $\{\vec{e}_{\mu\,\mathbf{x}}\}$, where μ labels the basis vector (e.g., $\mu = 0, 1, 2, 3$) and $\mathbf{x}$ labels the spacetime point.. Any four linearly independent basis vectors at each spacetime point will work; we do not not impose orthonormality or any other conditions in general, nor have we implied any relation to coordinates (although later we will). Given a basis, we may expand any vector field $\vec{A}$ as a linear combination of basis vectors:

$$\vec{A}_{\mathbf{x}} = A^{\mu}{}_{\mathbf{x}}\,\vec{e}_{\mu\,\mathbf{x}} = A^0{}_{\mathbf{x}}\vec{e}_{0\,\mathbf{x}} + A^1{}_{\mathbf{x}}\vec{e}_{1\,\mathbf{x}} + A^2{}_{\mathbf{x}}\vec{e}_{2\,\mathbf{x}} + A^3{}_{\mathbf{x}}\vec{e}_{3\,\mathbf{x}}\ . \tag{12}$$

Note our placement of subscripts and superscripts, chosen for consistency with the Einstein summation convention, which requires pairing one subscript with one superscript. The coefficients A^{μ} are called the components of the vector (often, the contravariant components). Note well that the coefficients A^{μ} depend on the basis vectors but $\vec{A}$ does not!

Similarly, we may choose a basis of one-form fields in which to expand one-forms like $\tilde{A}_{\mathbf{x}}$. Although any set of four linearly independent one-forms will suffice for each spacetime point, we prefer to choose a special one-form basis called the *dual basis* and denoted $\{\tilde{e}^{\mu}{}_{\mathbf{x}}\}$. Note that the placement of subscripts and superscripts is significant; we never use a subscript to label a basis one-form while we never use a superscript to label a basis vector. Therefore, $\tilde{e}^{\mu}$ is *not* related to $\vec{e}_{\mu}$ through the metric (eq. 9):

$\tilde{e}^\mu(\,\cdot\,) \neq \mathsf{g}(\vec{e}_\mu,\,\cdot\,)$. Rather, the dual basis one-forms are defined by imposing the following 16 requirements at each spacetime point:

$$\langle \tilde{e}^\mu{}_{\mathbf{x}}, \vec{e}_{\nu\,\mathbf{x}} \rangle = \delta^\mu{}_\nu \ , \tag{13}$$

where $\delta^\mu{}_\nu$ is the Kronecker delta, $\delta^\mu{}_\nu = 1$ if $\mu = \nu$ and $\delta^\mu{}_\nu = 0$ otherwise, with the same values for each spacetime point. (We must always distinguish subscripts from superscripts; the Kronecker delta always has one of each.) Equation (13) is a system of four linear equations at each spacetime point for each of the four quantities $\tilde{e}^\mu$ and it has a unique solution. (The reader may show that any nontrivial transformation of the dual basis one-forms will violate eq. 13.) Now we may expand any one-form field $\tilde{P}_{\mathbf{x}}$ in the basis of one-forms:

$$\tilde{P}_{\mathbf{x}} = P_{\mu\,\mathbf{x}}\,\tilde{e}^\mu{}_{\mathbf{x}} \ . \tag{14}$$

The component P_μ of the one-form $\tilde{P}$ is often called the covariant component to distinguish it from the contravariant component P^μ of the vector $\vec{P}$. In fact, because we have consistently treated vectors and one-forms as distinct, we should not think of these as being distinct "components" of the same entity at all.

There is a simple way to get the components of vectors and one-forms, using the fact that vectors are scalar functions of one-forms and vice versa. One simply evaluates the vector using the appropriate basis one-form:

$$\vec{A}(\tilde{e}^\mu) = \langle \tilde{e}^\mu, \vec{A} \rangle = \langle \tilde{e}^\mu, A^\nu \vec{e}_\nu \rangle = \langle \tilde{e}^\mu, \vec{e}_\nu \rangle A^\nu = \delta^\mu{}_\nu A^\nu = A^\mu \ , \tag{15}$$

and conversely for a one-form:

$$\tilde{P}(\vec{e}_\mu) = \langle \tilde{P}, \vec{e}_\mu \rangle = \langle P_\nu \tilde{e}^\nu, \vec{e}_\mu \rangle = \langle \tilde{e}^\nu, \vec{e}_\mu \rangle P_\nu = \delta^\nu{}_\mu P_\nu = P_\mu \ . \tag{16}$$

We have suppressed the spacetime point $\mathbf{x}$ for clarity, but it is always implied.

3.3 Tensor algebra

We can use the same ideas to expand tensors as products of components and basis tensors. First we note that a basis for a tensor of rank (m,n) is provided by the tensor product of m vectors and n one-forms. For example, a $(0,2)$ tensor like the metric tensor can be decomposed into basis tensors $\tilde{e}^\mu \otimes \tilde{e}^\nu$. The *components* of a tensor of rank (m,n), labeled with m superscripts and n subscripts, are obtained by evaluating the tensor using m basis one-forms and n basis vectors. For example, the components of the $(0,2)$ metric tensor, the $(2,0)$ inverse metric tensor and the $(1,1)$ identity tensor are

$$g_{\mu\nu} \equiv \mathsf{g}(\vec{e}_\mu, \vec{e}_\nu) = \vec{e}_\mu \cdot \vec{e}_\nu \ , \quad g^{\mu\nu} \equiv \mathsf{g}^{-1}(\tilde{e}^\mu, \tilde{e}^\nu) = \tilde{e}^\mu \cdot \tilde{e}^\nu \ , \quad \delta^\mu{}_\nu = \mathsf{I}(\tilde{e}^\mu, \vec{e}_\nu) = \langle \tilde{e}^\mu, \vec{e}_\nu \rangle \ . \tag{17}$$

(The last equation follows from eqs 4 and 13.) The tensors are given by summing over the tensor product of basis vectors and one-forms:

$$\mathsf{g} = g_{\mu\nu}\,\tilde{e}^{\mu} \otimes \tilde{e}^{\nu} \ , \quad \mathsf{g}^{-1} = g^{\mu\nu}\,\vec{e}_{\mu} \otimes \vec{e}_{\nu} \ , \quad \mathsf{I} = \delta^{\mu}{}_{\nu}\,\vec{e}_{\mu} \otimes \tilde{e}^{\nu} \ . \tag{18}$$

The reader should check that equation (18) follows from equations (17) and the duality condition equation (13).

Basis vectors and one-forms allow us to represent any tensor equations using components. For example, the dot product between two vectors or two one-forms and the scalar product between a one-form and a vector may be written using components as

$$\vec{A}\cdot\vec{B} = g_{\mu\nu}A^{\mu}A^{\nu} \ , \quad \langle\tilde{P},\vec{A}\rangle = P_{\mu}A^{\mu} \ , \quad \tilde{P}\cdot\tilde{Q} = g^{\mu\nu}P_{\mu}P_{\nu} \ . \tag{19}$$

The reader should prove these important results.

If two tensors of the same rank are equal in one basis, i.e., if all of their components are equal, then they are equal in any basis. While this mathematical result is obvious from the basis-free meaning of a tensor, it will have important physical implications in GR arising from the Equivalence Principle.

As we discussed above, the metric and inverse metric tensors allow us to transform vectors into one-forms and vice versa. If we evaluate the components of $\vec{V}$ and the one-form $\tilde{V}$ defined by equations (9) and (10), we get

$$V_{\mu} = \mathsf{g}(\vec{e}_{\mu},\vec{V}) = g_{\mu\nu}V^{\nu} \ , \quad V^{\mu} = \mathsf{g}^{-1}(\tilde{e}^{\mu},\tilde{V}) = g^{\mu\nu}V_{\nu} \ . \tag{20}$$

Because these two equations must hold for any vector $\vec{V}$, we conclude that the matrix defined by $g^{\mu\nu}$ is the inverse of the matrix defined by $g_{\mu\nu}$:

$$g^{\mu\kappa}g_{\kappa\nu} = \delta^{\mu}{}_{\nu} \ . \tag{21}$$

We also see that the metric and its inverse are used to lower and raise indices on components. Thus, given two vectors $\vec{V}$ and $\vec{W}$, we may evaluate the dot product any of four equivalent ways (cf. eq. 11):

$$\vec{V}\cdot\vec{W} = g_{\mu\nu}V^{\mu}W^{\nu} = V^{\mu}W_{\mu} = V_{\mu}W^{\mu} = g^{\mu\nu}V_{\mu}W_{\nu} \ . \tag{22}$$

In fact, the metric and its inverse may be used to transform tensors of rank (m,n) into tensors of any rank $(m+k, n-k)$ where $k = -m, -m+1, \ldots, n$. Consider, for example, a $(1,2)$ tensor T with components

$$T^{\mu}{}_{\nu\lambda} \equiv \mathsf{T}(\tilde{e}^{\mu},\vec{e}_{\nu},\vec{e}_{\lambda}) \ . \tag{23}$$

If we fail to plug in the one-form $\tilde{e}^{\mu}$, the result is the vector $T^{\kappa}{}_{\nu\lambda}\vec{e}_{\kappa}$. (A one-form must be inserted to return the quantity $T^{\kappa}{}_{\nu\lambda}$.) This vector may then be inserted into the metric tensor to give the components of a $(0,3)$ tensor:

$$T_{\mu\nu\lambda} \equiv \mathsf{g}(\vec{e}_{\mu}, T^{\kappa}{}_{\nu\lambda}\vec{e}_{\kappa}) = g_{\mu\kappa}T^{\kappa}{}_{\nu\lambda} \ . \tag{24}$$

We could now use the inverse metric to raise the third index, say, giving us the component of a $(1, 2)$ tensor distinct from equation (23):

$$T_{\mu\nu}{}^{\lambda} \equiv \mathsf{g}^{-1}(\tilde{e}^{\lambda}, T_{\mu\nu\kappa}\tilde{e}^{\kappa}) = g^{\lambda\kappa}T_{\mu\nu\kappa} = g^{\lambda\kappa}g_{\mu\rho}T^{\rho}{}_{\nu\kappa} \; . \tag{25}$$

In fact, there are 2^{m+n} different tensor spaces with ranks summing to $m+n$. The metric or inverse metric tensor allow all of these tensors to be transformed into each other.

Returning to equation (22), we see why we must distinguish vectors (with components V^{μ}) from one-forms (with components V_{μ}). The scalar product of two vectors requires the metric tensor while that of two one-forms requires the inverse metric tensor. In general, $g^{\mu\nu} \neq g_{\mu\nu}$. The *only* case in which the distinction is unnecessary is in flat (Lorentz) spacetime with orthonormal Cartesian basis vectors, in which case $g_{\mu\nu} = \eta_{\mu\nu}$ is everywhere the diagonal matrix with entries $(-1, +1, +1, +1)$. However, gravity curves spacetime. (Besides, we may wish to use curvilinear coordinates even in flat spacetime.) As a result, it is impossible to define a coordinate system for which $g^{\mu\nu} = g_{\mu\nu}$ everywhere. We must therefore distinguish vectors from one-forms and we must be careful about the placement of subscripts and superscripts on components.

At this stage it is useful to introduce a classification of vectors and one-forms drawn from special relativity with its Minkowski metric $\eta_{\mu\nu}$. Recall that a vector $\vec{A} = A^{\mu}\vec{e}_{\mu}$ is called spacelike, timelike or null according to whether $\vec{A} \cdot \vec{A} = \eta_{\mu\nu}A^{\mu}A^{\nu}$ is positive, negative or zero, respectively. In a Euclidean space, with positive definite metric, $\vec{A} \cdot \vec{A}$ is never negative. However, in the Lorentzian spacetime geometry of special relativity, time enters the metric with opposite sign so that it is possible to have $\vec{A} \cdot \vec{A} < 0$. In particular, the four-velocity $u^{\mu} = dx^{\mu}/d\tau$ of a massive particle (where $d\tau$ is proper time) is a timelike vector. This is seen most simply by performing a Lorentz boost to the rest frame of the particle in which case $u^t = 1$, $u^x = u^y = u^z = 0$ and $\eta_{\mu\nu}u^{\mu}u^{\nu} = -1$. Now, $\eta_{\mu\nu}u^{\mu}u^{\nu}$ is a Lorentz scalar so that $\vec{u} \cdot \vec{u} = -1$ in any Lorentz frame. Often this is written $\vec{p} \cdot \vec{p} = -m^2$ where $p^{\mu} = mu^{\mu}$ is the four-momentum for a particle of mass m. For a massless particle (e.g., a photon) the proper time vanishes but the four-momentum is still well-defined with $\vec{p} \cdot \vec{p} = 0$: the momentum vector is null. We adopt the same notation in general relativity, replacing the Minkowski metric (components $\eta_{\mu\nu}$) with the actual metric g and evaluating the dot product using $\vec{A} \cdot \vec{A} = \mathsf{g}(\vec{A}, \vec{A}) = g_{\mu\nu}A^{\mu}A^{\nu}$. The same classification scheme extends to one-forms using g^{-1}: a one-form $\tilde{P}$ is spacelike, timelike or null according to whether $\tilde{P} \cdot \tilde{P} = \mathsf{g}^{-1}(\tilde{P}, \tilde{P}) = g^{\mu\nu}P_{\mu}P_{\nu}$ is positive, negative or zero, respectively. Finally, a vector is called a unit vector if $\vec{A} \cdot \vec{A} = \pm 1$ and similarly for a one-form. The four-velocity of a massive particle is a timelike unit vector.

Now that we have introduced basis vectors and one-forms, we can define the contraction of a tensor. Contraction pairs two arguments of a rank (m, n) tensor: one vector and one one-form. The arguments are replaced by basis vectors and one-forms and summed over. For example, consider the $(1, 3)$ tensor R, which may be contracted on its second vector argument to give a $(0,2)$ tensor also denoted R but distinguished by its shorter argument list:

$$\mathsf{R}(\vec{A},\vec{B}) = \delta^\lambda{}_\kappa\,\mathsf{R}(\tilde{e}^\kappa,\vec{A},\vec{e}_\lambda,\vec{B}) = \sum_{\lambda=0}^{3}\mathsf{R}(\tilde{e}^\lambda,\vec{A},\vec{e}_\lambda,\vec{B}) \ . \tag{26}$$

In Section 6 we will define the Riemann curvature tensor of rank $(1,3)$; its contraction, defined by equation (26), is called the Ricci tensor. Although the contracted tensor would appear to depend on the choice of basis because its definition involves the basis vectors and one-forms, the reader may show that it is actually invariant under a change of basis (and is therefore a tensor) as long as we use dual one-form and vector bases satisfying equation (13). Equation (26) becomes somewhat clearer if we express it entirely using tensor components:

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu} \ . \tag{27}$$

Summation over λ is implied. Contraction may be performed on any pair of covariant and contravariant indices; different tensors result.

3.4 Change of basis

We have made no restrictions upon our choice of basis vectors $\vec{e}_\mu$. Our basis vectors are simply a linearly independent set of four vector fields allowing us to express any vector as a linear combination of basis vectors. The choice of basis is not unique; we may transform from one basis to another simply by defining four linearly independent combinations of our original basis vectors. Given one basis $\{\vec{e}_\mu\,\mathbf{x}\}$, we may define another basis $\{\vec{e}_{\mu'}\,\mathbf{x}\}$, distinguished by placing a prime on the labels, as follows:

$$\vec{e}_{\mu'}\,\mathbf{x} = \Lambda^\nu{}_{\mu'}\,\mathbf{x}\,\vec{e}_\nu\,\mathbf{x} \ . \tag{28}$$

The prime is placed on the index rather than on the vector only for notational convenience; we do not imply that the new basis vectors are a permutation of the old ones.

Any linearly independent linear combination of old basis vectors may be selected as the new basis vectors. That is, any nonsingular four-by-four matrix may be used for $\Lambda^\nu{}_{\mu'}$. The transformation is not restricted to being a Lorentz transformation; the local reference frames defined by the bases are not required to be inertial (or, in the presence of gravity, freely-falling). Because the transformation matrix is assumed to be nonsingular, the transformation may be inverted:

$$\vec{e}_\nu\,\mathbf{x} = \Lambda^{\mu'}{}_\nu\,\mathbf{x}\,\vec{e}_{\mu'}\,\mathbf{x} \ , \quad \Lambda^\mu{}_{\kappa'}\Lambda^{\kappa'}{}_\nu \equiv \ \delta^\mu{}_\nu \ . \tag{29}$$

Comparing equations (28) and (29), note that in our notation the inverse matrix places the prime on the other index. Primed indices are never summed together with unprimed indices.

If we change basis vectors, we must also transform the basis one-forms so as to preserve the duality condition equation (13). The reader may verify that, given the transformations of equations (28) and (29), the new dual basis one-forms are

$$\tilde{e}^{\mu'}{}_{\mathbf{x}} = \Lambda^{\mu'}{}_{\nu\,\mathbf{x}}\tilde{e}^{\nu}{}_{\mathbf{x}} \ . \tag{30}$$

We may also write the transformation matrix and its inverse by scalar products of the old and new basis vectors and one-forms (dropping the subscript $\mathbf{x}$ for clarity):

$$\Lambda^{\nu}{}_{\mu'} = \langle \tilde{e}^{\nu}, \vec{e}_{\mu'} \rangle \ , \quad \Lambda^{\mu'}{}_{\nu} = \langle \tilde{e}^{\mu'}, \vec{e}_{\nu} \rangle \ . \tag{31}$$

Apart from the basis vectors and one-forms, a vector $\vec{A}$ and a one-form $\tilde{P}$ are, by definition, invariant under a change of basis. Their components are not. For example, using equation (29) or (31) we find

$$\vec{A} = A^{\nu}\vec{e}_{\nu} = A^{\mu'}\vec{e}_{\mu'} \ , \quad A^{\mu'} = \langle \tilde{e}^{\mu'}, \vec{A} \rangle = \Lambda^{\mu'}{}_{\nu}A^{\nu} \ . \tag{32}$$

The vector components transform oppositely to the basis vectors (eq. 28). One-form components transform like basis vectors, as suggested by the fact that both are labeled with a subscript:

$$\tilde{A} = A_{\nu}\tilde{e}^{\nu} = A_{\mu'}\tilde{e}^{\mu'} \ , \quad A_{\mu'} = \langle \tilde{A}, \vec{e}_{\mu'} \rangle = \Lambda^{\nu}{}_{\mu'}A_{\nu} \ . \tag{33}$$

Note that if the components of two vectors or two one-forms are equal in one basis, they are equal in any basis.

Tensor components also transform under a change of basis. The new components may be found by recalling that a (m, n) tensor is a function of m vectors and n one-forms and that its components are gotten by evaluating the tensor using the basis vectors and one-forms (e.g., eq. 17). For example, the metric components are transformed under the change of basis of equation (28) to

$$g_{\mu'\nu'} \equiv \mathsf{g}(\vec{e}_{\mu'}, \vec{e}_{\nu'}) = g_{\alpha\beta}\,\tilde{e}^{\alpha}(\vec{e}_{\mu'})\tilde{e}^{\beta}(\vec{e}_{\nu'}) = g_{\alpha\beta}\,\Lambda^{\alpha}{}_{\mu'}\Lambda^{\beta}{}_{\nu'} \ . \tag{34}$$

(Recall that "evaluating" a one-form or vector means using the scalar product, eq. 1.) We see that the covariant components of the metric (i.e., the lower indices) transform exactly like one-form components. Not surprisingly, the components of a tensor of rank (m, n) transform like the product of m vector components and n one-form components. If the components of two tensors of the same rank are equal in one basis, they are equal in any basis.

3.5 Coordinate bases

We have made no restrictions upon our choice of basis vectors $\vec{e}_\mu$. Before concluding our formal introduction to tensors, we introduce one more idea: a coordinate system. A coordinate system is simply a set of four differentiable *scalar* fields $x^\mu{}_{\mathbf{x}}$ (*not* one *vector* field — note that μ labels the coordinates and not vector components) that attach a unique set of labels to each spacetime point $\mathbf{x}$. That is, no two points are allowed to have identical values of all four scalar fields and the coordinates must vary smoothly throughout spacetime (although we will tolerate occasional flaws like the coordinate singularities at $r = 0$ and $\theta = 0$ in spherical polar coordinates). Note that we impose no other restrictions on the coordinates. The freedom to choose different coordinate systems is available to us even in a Euclidean space; there is nothing sacred about Cartesian coordinates. This is even more true in a non-Euclidean space, where Cartesian coordinates covering the whole space do not exist.

Coordinate systems are useful for three reasons. First and most obvious, they allow us to label each spacetime point by a set of numbers (x^0, x^1, x^2, x^3). The second and more important use is in providing a special set of basis vectors called a coordinate basis. Suppose that two infinitesimally close spacetime points have coordinates x^μ and $x^\mu + dx^\mu$. The infinitesimal difference vector between the two points, denoted $d\vec{x}$, is a vector defined at $\mathbf{x}$. We define the coordinate basis as the set of four basis vectors $\vec{e}_{\mu\,\mathbf{x}}$ such that the components of $d\vec{x}$ are dx^μ:

$$d\vec{x} \equiv dx^\mu \vec{e}_\mu \quad \text{defines } \vec{e}_\mu \text{ in a coordinate basis .} \tag{35}$$

From the trivial appearance of this equation the reader may incorrectly think that we have imposed no constraints on the basis vectors. However, that is not so. According to equation (35), the basis vector $\vec{e}_{0\,\mathbf{x}}$, for example, must point in the direction of increasing x^0 at point $\mathbf{x}$. This corresponds to a unique direction in four-dimensional spacetime just as the direction of increasing latitude corresponds to a unique direction (north) at a given point on the earth. In more mathematical treatments (e.g. Walk 1984), $\vec{e}_\mu$ is associated with the directional derivative $\partial/\partial x^\mu$ at $\mathbf{x}$.

Note that not all bases are coordinate bases. If we wanted to be perverse we could define a non-coordinate basis by, for example, permuting the labels on the basis vectors but not those on the coordinates (which, after all, are not the components of a vector). In this case $\langle \tilde{e}^\mu, d\vec{x} \rangle$, the component of $d\vec{x}$ for basis vector $\vec{e}_\mu$, would not equal the coordinate differential dx^μ. This would violate nothing we have written so far except equation (35). Later we will discover more natural ways that non-coordinate bases may arise.

The coordinate basis $\{\vec{e}_\mu\}$ defined by equation (35) has a dual basis of one-forms $\{\tilde{e}^\mu\}$ defined by equation (13). The dual basis of one-forms is related to the gradient. We obtain this relation as follows. Consider any scalar field $f_{\mathbf{x}}$. Treating f as a function of the coordinates, the difference in f between two infinitesimally close points is

$$df = \frac{\partial f}{\partial x^\mu}\, dx^\mu \equiv \partial_\mu f\, dx^\mu \ . \tag{36}$$

Equation (36) may be taken as the definition of the components of the gradient (with an alternative brief notation for the partial derivative). However, partial derivatives depend on the coordinates, while the gradient (covariant derivative) should not. What, then, is the gradient — is it a vector or a one-form?

From equation (36), because df is a scalar and dx^μ is a vector component, $\partial f/\partial x^\mu$ must be the component of a one-form, not a vector. The notation ∂_μ, with its covariant (subscript) index, reinforces our view that the partial derivative is the component of a one-form and not a vector. We denote the gradient one-form by $\tilde{\nabla}$. Like all one-forms,

The gradient may be decomposed into a sum over basis one-forms $\tilde{e}^\mu$. Using equation (36) and the equation (13) as the requirement for a dual basis, we conclude that the gradient is

$$\tilde{\nabla} \equiv \tilde{e}^\mu\, \partial_\mu \quad \text{in a coordinate basis} \ . \tag{37}$$

Note that we must write the basis one-form to the left of the partial derivative operator, for the basis one-form itself may depend on position! We will return to this point in Section 4 when we discuss the covariant derivative. In the present case, it is clear from equation (36) that we must let the derivative act only on the function f. We can now rewrite equation (36) in the coordinate-free manner

$$df = \langle \tilde{\nabla} f, d\vec{x} \rangle \ . \tag{38}$$

If we want the directional derivative of f along any particular direction, we simply replace $d\vec{x}$ by a vector pointing in the desired direction (e.g., the tangent vector to some curve). Also, if we let $f_{\mathbf{X}}$ equal one of the coordinates, using equation (37) the gradient gives us the corresponding basis one-form:

$$\tilde{\nabla} x^\mu = \tilde{e}^\mu \quad \text{in a coordinate basis} \ . \tag{39}$$

The third use of coordinates is that they can be used to describe the distance between two points of spacetime. However, coordinates alone are not enough. We also need the metric tensor. We write the squared distance between two spacetime points as

$$ds^2 = |d\vec{x}\,|^2 \equiv \mathsf{g}(d\vec{x}, d\vec{x}) = d\vec{x} \cdot d\vec{x} \ . \tag{40}$$

IThis equation, true in any basis because it is a scalar equation that makes no reference to components, is taken as the *definition* of the metric tensor. Up to now the metric could have been any symmetric $(0,2)$ tensor. But, if we insist on being able to measure distances, given an infinitesimal difference vector $d\vec{x}$, only one $(0,2)$ tensor can give the squared distance. We define the metric tensor to be that tensor. Indeed, the squared magnitude of *any* vector $\vec{A}$ is $|\vec{A}|^2 \equiv \mathsf{g}(\vec{A}, \vec{A})$.

Now we specialize to a coordinate basis, using equation (35) for $d\vec{x}$. In a coordinate basis (and *only* in a coordinate basis), the squared distance is called the line element and takes the form

$$ds^2 = g_{\mu\nu\,\mathbf{X}} dx^\mu dx^\nu \quad \text{in a coordinate basis .} \tag{41}$$

We have used equation (17) to get the metric components.

If we transform coordinates, we will have to change our vector and one-form bases. Suppose that we transform from $x^\mu{}_{\mathbf{X}}$ to $x^{\mu'}{}_{\mathbf{X}}$, with a prime indicating the new coordinates. For example, in the Euclidean plane we could transform from Cartesian coordinate ($x^1 = x, x^2 = y$) to polar coordinates ($x^{1'} = r, x^{2'} = \theta$): $x = r\cos\theta$, $y = r\sin\theta$. A one-to-one mapping is given from the old to new coordinates, allowing us to define the Jacobian matrix $\Lambda^{\mu'}{}_\nu \equiv \partial x^{\mu'}/\partial x^\nu$ and its inverse $\Lambda^\nu{}_{\mu'} = \partial x^\nu/\partial x^{\mu'}$. Vector components transform like $dx^{\mu'} = (\partial x^{\mu'}/\partial x^\nu)\, dx^\nu$. Transforming the basis vectors, basis one-forms, and tensor components is straightforward using equations (28)–(34). The reader should verify that equations (35), (37), (39) and (41) remain valid after a coordinate transformation.

We have now introduced many of the basic ingredients of tensor algebra that we will need in general relativity. Before moving on to more advanced concepts, let us reflect on our treatment of vectors, one-forms and tensors. The mathematics and notation, while straightforward, are complicated. Can we simplify the notation without sacrificing rigor?

One way to modify our notation would be to abandon ths basis vectors and one-forms and to work only with components of tensors. We could have defined vectors, one-forms and tensors from the outset in terms of the transformation properties of their components. However, the reader should appreciate the clarity of the geometrical approach that we have adopted. Our notation has forced us to distinguish physical objects like vectors from basis-dependent ones like vector components. As long as the definition of a tensor is not forgotten, computations are straightforward and unambiguous. Moreover, adopting a basis did not force us to abandon geometrical concepts. On the contrary, computations are made easier and clearer by retaining the notation and meaning of basis vectors and one-forms.

3.6 Isomorphism of vectors and one-forms

Although vectors and one-forms are distinct objects, there is a strong relationship between them. In fact, the linear space of vectors is isomorphic to the dual vector space of one-forms (Wald 1984). Every equation or operation in one space has an equivalent equation or operation in the other space. This isomorphism can be used to hide the distinction between one-forms and vectors in a way that simplifies the notation. This approach is unusual (I haven't seen it published anywhere) and is not recommended in formal work but it may be pedagogically useful.

As we saw in equations (9) and (10), the link between the vector and dual vector spaces is provided by g and g^{-1}. If $\vec{A} = \vec{B}$ (components $A^\mu = B^\mu$), then $\tilde{A} = \tilde{B}$ (components $A_\mu = B_\mu$) where $A_\mu = g_{\mu\nu} A^\nu$ and $B_\mu = g_{\mu\nu} B^\nu$. So, why do we bother with one-forms when vectors are sufficient? The answer is that tensors may be functions of both one-forms and vectors. However, there is also an isomorphism among tensors of different rank. We have just argued that the tensor spaces of rank $(1, 0)$ (vectors) and $(0, 1)$ are isomorphic. In fact, all 2^{m+n} tensor spaces of rank (m, n) with fixed $m + n$ are isomorphic. The metric and inverse metric tensors link together these spaces, as exemplified by equations (24) and (25).

The isomorphism of different tensor spaces allows us to introduce a notation that unifies them. We could effect such a unification by discarding basis vectors and one-forms and working only with components, using the components of the metric tensor and its inverse to relate components of different types of tensors as in equations (24) and (25). However, this would require sacrificing the coordinate-free geometrical interpretation of vectors. Instead, we introduce a notation that replaces one-forms with vectors and (m, n) tensors with $(m + n, 0)$ tensors in general. We do this by replacing the basis one-forms $\tilde{e}^\mu$ with a set of *vectors* defined as in equation (10):

$$\vec{e}^{\,\mu}(\,\cdot\,) \equiv \mathsf{g}^{-1}(\tilde{e}^\mu, \,\cdot\,) = g^{\mu\nu} \vec{e}_\nu(\,\cdot\,) \ . \tag{42}$$

We will refer to $\vec{e}^{\,\mu}$ as a *dual basis vector* to contrast it from both the basis vector $\vec{e}_\mu$ and the basis one-form $\tilde{e}^\mu$. The dots are present in equation (42) to remind us that a one-form may be inserted to give a scalar. However, we no longer need to use one-forms. Using equation (42), given the components A_μ of any one-form $\tilde{A}$, we may form the vector $\vec{A}$ defined by equation (10) as follows:

$$\vec{A} = A_\mu \vec{e}^{\,\mu} = A_\mu g^{\mu\nu} \vec{e}_\nu = A^\mu \vec{e}_\mu \ . \tag{43}$$

The reader should verify that $\vec{A} = A_\mu \vec{e}^{\,\mu}$ is invariant under a change of basis because $\vec{e}^{\,\mu}$ transforms like a basis one-form.

The isomorphism of one-forms and vectors means that we can replace all one-forms with vectors in any tensor equation. Tildes may be replaced with arrows. The scalar product between a one-form and a vector is replaced by the dot product using the metric (eq. 10 or 42). The only rule is that we must treat a dual basis vector with an upper index like a basis one-form:

$$\vec{e}_\mu \cdot \vec{e}_\nu = g_{\mu\nu} \ , \quad \vec{e}^{\,\mu} \cdot \vec{e}_\nu = \langle \tilde{e}^{\,\mu}, \vec{e}_\nu \rangle = \delta^\mu{}_\nu \ , \quad \vec{e}^{\,\mu} \cdot \vec{e}^{\,\nu} = \tilde{e}^{\,\mu} \cdot \tilde{e}^{\,\nu} = g^{\mu\nu} \ . \tag{44}$$

The reader should verify equations (44) using equations (17) and (42). Now, if we need the contravariant component of a vector, we can get it from the dot product with the dual basis vector instead of from the scalar product with the basis one-form:

$$A^\mu = \vec{e}^{\,\mu} \cdot \vec{A} = \langle \tilde{e}^\mu, \vec{A} \rangle \ . \tag{45}$$

We may also apply this recipe to convert the gradient one-form $\tilde{\nabla}$ (eq. 37) to a vector, though we must not allow the dual basis vector to be differentiated:

$$\vec{\nabla} = \vec{e}^{\,\mu} \partial_\mu = g^{\mu\nu} \vec{e}_\mu \, \partial_\nu \quad \text{in a coordinate basis} \ . \tag{46}$$

It follows at once that the dual basis vector (in a coordinate basis) is the vector gradient of the coordinate: $\vec{e}^{\,\mu} = \vec{\nabla} x^\mu$. This equation is isomorphic to equation (39).

The basis vectors and dual basis vectors, through their tensor products, also give a basis for higher-rank tensors. Again, the rule is to replace the basis one-forms with the corresponding dual basis vectors. Thus, for example, we may write the rank $(2,0)$ metric tensor in any of four ways:

$$\mathsf{g} = g_{\mu\nu} \vec{e}^{\,\mu} \otimes \vec{e}^{\,\nu} = g^\mu{}_\nu \vec{e}_\mu \otimes \vec{e}^{\,\nu} = g_\mu{}^\nu \vec{e}^{\,\mu} \otimes \vec{e}_\nu = g^{\mu\nu} \vec{e}_\mu \otimes \vec{e}_\nu \ . \tag{47}$$

In fact, by comparing this with equation (18) the reader will see that what we have written is actually the *inverse* metric tensor g^{-1}, which is isomorphic to g through the replacement of $\tilde{e}^\mu$ with $\vec{e}^{\,\mu}$. But, what are the mixed components of the metric, $g^\mu{}_\nu$ and $g_\mu{}^\nu$? From equations (13) and (42), we see that they both equal the Kronecker delta $\delta^\mu{}_\nu$. Consequently, the metric tensor is isomorphic to the identity tensor as well as to its inverse! However, this is no miracle; it was guaranteed by our definition of the dual basis vectors and by the fact we defined g^{-1} to invert the mapping from vectors to one-forms implied by g. The reader may fear that we have defined away the metric by showing it to be isomorphic to the identity tensor. However, this is not the case. We need the metric tensor components to obtain $\vec{e}^{\,\mu}$ from $\vec{e}_\mu$ or A^μ from A_μ. We cannot take advantage of the isomorphism of different tensor spaces without the metric. Moreover, as we showed in equation (40), the metric plays a fundamental role in giving the squared magnitude of a vector. In fact, as we will see later, the metric contains all of the information about the geometrical properties of spacetime. Clearly, the metric must play a fundamental role in general relativity.

3.7 Example: Euclidean plane

We close this section by applying tensor concepts to a simple example: the Euclidean plane. This flat two-dimensional space can be covered by Cartesian coordinates (x,y) with line element and metric components

$$ds^2 = dx^2 + dy^2 \quad \Rightarrow \quad g_{xx} = g_{yy} = 1 \ , \ g_{xy} = g_{yx} = 0 \ . \tag{48}$$

We prefer to use the coordinate names themselves as component labels rather than using numbers (e.g. g_{xx} rather than g_{11}). The basis vectors are denoted $\vec{e}_x$ and $\vec{e}_y$, and their use in plane geometry and linear algebra is standard. Basis one-forms appear unnecessary because the metric tensor is just the identity tensor in this basis. Consequently the dual basis vectors (eq. 42) are $\vec{e}^{\,x} = \vec{e}_x$, $\vec{e}^{\,y} = \vec{e}_y$ and no distinction is needed between superscripts and subscripts.

However, there is nothing sacred about Cartesian coordinates. Consider polar coordinates (ρ, θ), defined by the transformation $x = \rho\cos\theta$, $y = \rho\sin\theta$. A simple exercise in partial derivatives yields the line element in polar coordinates:

$$ds^2 = d\rho^2 + \rho^2 d\theta^2 \quad \Rightarrow \quad g_{\rho\rho} = 1 \ , \ g_{\theta\theta} = \rho^2 \ , \ g_{\rho\theta} = g_{\theta\rho} = 0 \ . \tag{49}$$

This appears eminently reasonable until, perhaps, one considers the basis vectors $\vec{e}_\rho$ and $\vec{e}_\theta$, recalling that $g_{\mu\nu} = \vec{e}_\mu \cdot \vec{e}_\nu$. Then, while $\vec{e}_\rho \cdot \vec{e}_\rho = 1$ and $\vec{e}_\rho \cdot \vec{e}_\theta = 0$, $\vec{e}_\theta \cdot \vec{e}_\theta = \rho^2$: $\vec{e}_\theta$ is not a unit vector! The new basis vectors are easily found in terms of the Cartesian basis vectors and components using equation (28):

$$\vec{e}_\rho = \frac{x}{\sqrt{x^2+y^2}}\,\vec{e}_x + \frac{y}{\sqrt{x^2+y^2}}\,\vec{e}_y \ , \quad \vec{e}_\theta = -y\,\vec{e}_x + x\,\vec{e}_y \ . \tag{50}$$

The polar *unit* vectors are $\hat{\rho} = \vec{e}_\rho$ and $\hat{\theta} = \rho^{-1}\vec{e}_\theta$.

Why does our formalism give us non-unit vectors? The answer is because we insisted that our basis vectors be a *coordinate* basis (eqs. 35, 37, 39 and 41). In terms of the orthonormal unit vectors, the difference vector between points (ρ, θ) and $(\rho+d\rho, \theta+d\theta)$ is $d\vec{x} = \hat{\rho}\,d\rho + \hat{\theta}\,\rho d\theta$. In the coordinate basis this takes the simpler form $d\vec{x} = \vec{e}_\rho\,d\rho + \vec{e}_\theta\,d\theta = dx^\mu \vec{e}_\mu$. In the coordinate basis we don't have to worry about normalizing our vectors; all information about lengths is carried instead by the metric. In the *non*-coordinate basis of orthonormal vectors $\{\hat{\rho}, \hat{\theta}\}$ we have to make a separate note that the distance elements are $d\rho$ and $\rho d\theta$.

In the non-coordinate basis we can no longer use equation (41) for the line element. We must instead use equation (40). The metric components in the non-coordinate basis $\{\hat{\rho}, \hat{\theta}\}$ are

$$g_{\hat{\rho}\hat{\rho}} = g_{\hat{\theta}\hat{\theta}} = 1 \ , \ g_{\hat{\rho}\hat{\theta}} = g_{\hat{\theta}\hat{\rho}} = 0 \ . \tag{51}$$

The reader may also verify this result by transforming the components of the metric from the basis $\{\vec{e}_\rho, \vec{e}_\theta\}$ to $\{\hat{\rho}, \hat{\theta}\}$ using equation (34) with $\Lambda^\rho{}_{\hat{\rho}} = 1$, $\Lambda^\theta{}_{\hat{\theta}} = \rho^{-1}$. Now, equation (40) still gives the distance squared, but we are responsible for remembering $d\vec{x} = \hat{\rho}\,d\rho + \hat{\theta}\,\rho d\theta$. In a non-coordinate basis, the metric will not tell us how to measure distances in terms of coordinate differentials.

With a non-coordinate basis, we must sacrifice equations (35) and (41). Nonetheless, for some applications it proves convenient to introduce an orthonormal non-coordinate basis called a tetrad basis. Tetrads are discussed by Wald (1984) and Misner et al (1973).

The use of non-coordinate bases also complicates the gradient (eqs. 37, 39 and 46). In our polar coordinate basis (eq. 49), the inverse metric components are

$$g^{\rho\rho} = 1 \; , \; g^{\theta\theta} = \rho^{-2} \; , \; g^{\rho\theta} = g^{\theta\rho} = 0 \; . \tag{52}$$

(The matrix $g_{\mu\nu}$ is diagonal, so its inverse is also diagonal with entries given by the reciprocals.) The basis one-forms obey the rules $\tilde{e}^\mu \cdot \tilde{e}^\nu = g^{\mu\nu}$. They are isomorphic to the dual basis vectors $\vec{e}^{\,\mu} = g^{\mu\nu}\vec{e}_\nu$ (eq. 42). Thus, $\vec{e}^{\,\rho} = \vec{e}_\rho = \hat{\rho}$, $\vec{e}^{\,\theta} = \rho^{-2}\vec{e}_\theta = \rho^{-1}\hat{\theta}$. Equation (37) gives the gradient one-form as $\tilde{\nabla} = \tilde{e}^\rho \, (\partial/\partial\rho) + \tilde{e}^\theta \, (\partial/\partial\theta)$. Expressing this as a vector (eq. 46) we get

$$\vec{\nabla} = \vec{e}^{\,\rho} \frac{\partial}{\partial\rho} + \vec{e}^{\,\theta} \frac{\partial}{\partial\theta} = \hat{\rho} \, \frac{\partial}{\partial\rho} + \hat{\theta} \, \frac{1}{\rho} \frac{\partial}{\partial\theta} \; . \tag{53}$$

The gradient is simpler in the coordinate basis. The coordinate basis has the added advantage that we can get the dual basis vectors (or the basis one-forms) by applying the gradient to the coordinates (eq. 46): $\vec{e}^{\,\rho} = \vec{\nabla}\rho$, $\vec{e}^{\,\theta} = \vec{\nabla}\theta$.

From now on, unless otherwise noted, we will assume that our basis vectors are a coordinate basis. We will use one-forms and vectors interchangeably through the mapping provided by the metric and inverse metric (eqs. 9, 10 and 42). Readers who dislike one-forms may convert the tildes to arrows and use equations (44) to obtain scalars from scalar products and dot products.

4 Differentiation and Integration

In this section we discuss differentiation and integration in curved spacetime. These might seem like a delicate subjects but, given the tensor algebra that we have developed, tensor calculus is straightforward.

4.1 Gradient of a scalar

Consider first the gradient of a scalar field $f_{\mathbf{X}}$. We have already shown in Section 2 that the gradient operator $\tilde{\nabla}$ is a one-form (an object that is invariant under coordinate transformations) and that, in a coordinate basis, its components are simply the partial derivatives with respect to the coordinates:

$$\tilde{\nabla} f = (\partial_\mu f) \, \tilde{e}^\mu = (\nabla_\mu f) \, \tilde{e}^\mu \; , \tag{54}$$

where $\partial_\mu \equiv (\partial/\partial x^\mu)$. We have introduced a second notation, ∇_μ, called the *covariant derivative* with respect to x^μ. By definition, the covariant derivative behaves like the component of a one-form. But, from equation (54), this is also true of the partial derivative operator ∂_μ. Why have we introduced a new symbol?

Before answering this question, let us first note that the gradient, because it behaves like a tensor of rank $(0,1)$ (a one-form), changes the rank of a tensor field from (m,n) to $(m,n+1)$. (This is obviously true for the gradient of a scalar field, with $m=n=0$.) That is, application of the gradient is like taking the tensor product with a one-form. The difference is that the components are not the product of the components, because ∇_μ is not a number. Nevertheless, the resulting object must be a tensor of rank $(m,n+1)$; e.g., its components must transform like components of a $(m,n+1)$ tensor. The gradient of a scalar field f is a $(0,1)$ tensor with components $(\partial_\mu f)$.

4.2 Gradient of a vector: covariant derivative

The reason that we have introduced a new symbol for the derivative will become clear when we take the gradient of a vector field $\vec{A}_{\mathbf{x}} = A^\mu{}_{\mathbf{x}}\vec{e}_{\mu\,\mathbf{x}}$. In general, the basis vectors are functions of position as are the vector components! So, the gradient must act on both. In a coordinate basis, we have

$$\tilde{\nabla}\vec{A} = \tilde{\nabla}(A^\nu\vec{e}_\nu) = \tilde{e}^\mu\partial_\mu(A^\nu\vec{e}_\nu) = (\partial_\mu A^\nu)\,\tilde{e}^\mu\vec{e}_\nu + A^\nu\tilde{e}^\mu\,(\partial_\mu\vec{e}_\nu) \equiv (\nabla_\mu A^\nu)\,\tilde{e}^\mu\vec{e}_\nu\ . \tag{55}$$

We have dropped the tensor product symbol $\otimes$ for notational convenience although it is still implied. Note that we must be careful to preserve the ordering of the vectors and tensors and we must not confuse subscripts and superscripts. Otherwise, taking the gradient of a vector is straightforward. The result is a $(1,1)$ tensor with components $\nabla_\mu A^\nu$. But now $\nabla_\mu \neq \partial_\mu$! This is why we have introduced a new derivative symbol. We reserve the covariant derivative notation ∇_μ for the actual components of the gradient of a tensor. We note that the alternative notation $A^\nu{}_{;\mu} = \nabla_\mu A^\nu$ is often used, replacing the comma of a partial derivative $A^\nu{}_{,\mu} = \partial_\mu A^\nu$ with a semicolon for the covariant derivative. The difference seems mysterious only when we ignore basis vectors and stick entirely to components. As equation (55) shows, vector notation makes it clear why there is a difference.

Equation (55) by itself does not help us evaluate the gradient of a vector because we do not yet know what the gradients of the basis vectors are. However, they are straightforward to determine in a coordinate basis. First we note that, geometrically, $\partial_\mu\vec{e}_\nu$ is a vector at $\mathbf{x}$: it is the difference of two vectors at infinitesimally close points, divided by a coordinate interval. (The easiest way to tell that $\partial_\mu\vec{e}_\nu$ is a vector is to note that it has one arrow!) So, like all vectors, it must be a linear combination of basis vectors at $\mathbf{x}$. We can write the most general possible linear combination as

$$\partial_\mu\vec{e}_{\nu\,\mathbf{x}} \equiv \Gamma^\lambda{}_{\mu\nu\,\mathbf{x}}\,\vec{e}_{\lambda\,\mathbf{x}}\ . \tag{56}$$

4.3 Christoffel symbols

We have introduced in equation (56) a set of coefficients, ${\Gamma^\lambda}_{\mu\nu}$, called the connection coefficients or *Christoffel symbols*. (Technically, the term Christoffel symbols is reserved for a coordinate basis.) It should be noted at the outset that, despite their appearance, the Christoffel symbols are not the components of a $(1,2)$ tensor. Rather, they may be considered as a set of four $(1,1)$ tensors, one for each basis vector $\vec{e}_\nu$, because $\tilde{\nabla}\vec{e}_\nu = {\Gamma^\lambda}_{\mu\nu}\tilde{e}^\mu\vec{e}_\lambda$. However, it is not useful to think of the Christoffel symbols as tensor components for fixed ν because, under a change of basis, the basis vectors $\vec{e}_\nu$ themselves change and therefore the four $(1,1)$ tensors must also change. So, forget about the Christoffel symbols defining a tensor. They are simply a set of coefficients telling us how to differentiate basis vectors. Whatever their values, the components of the gradient of $\vec{A}$, known also as the covariant derivative of A^ν, are, from equations (55) and (56),

$$\nabla_\mu A^\nu = \partial_\mu A^\nu + {\Gamma^\nu}_{\mu\lambda}A^\lambda \ . \tag{57}$$

How does one determine the values of the Christoffel symbols? That is, how does one evaluate the gradients of the basis vectors? One way is to express the basis vectors in terms of another set whose gradients are known. For example, consider polar coordinates (ρ,θ) in the Cartesian plane as discussed in Section 2. The polar coordinate basis vectors were given in terms of the Cartesian basis vectors in equation (50). We know that the gradients of the Cartesian basis vectors vanish and we know how to transform from Cartesian to polar coordinates. It is a straightforward and instructive exercise from this to compute the gradients of the polar basis vectors:

$$\tilde{\nabla}\vec{e}_\rho = \frac{1}{\rho}\,\tilde{e}^\theta\otimes\vec{e}_\theta \ , \quad \tilde{\nabla}\vec{e}_\theta = \frac{1}{\rho}\,\tilde{e}^\rho\otimes\vec{e}_\theta - \rho\,\tilde{e}^\theta\otimes\vec{e}_\rho \ . \tag{58}$$

(We have restored the tensor product symbol as a reminder of the tensor nature of the objects in eq. 58.) From equations (56) and (58) we conclude that the nonvanishing Christoffel symbols are

$${\Gamma^\theta}_{\theta\rho} = {\Gamma^\theta}_{\rho\theta} = \rho^{-1} \ , \quad {\Gamma^\rho}_{\theta\theta} = -\rho \ . \tag{59}$$

It is instructive to extend this example further. Suppose that we add the third dimension, with coordinate z, to get a three-dimensional Euclidean space with cylindrical coordinates (ρ,θ,z). The line element (cf. eq. 49) now becomes $ds^2 = d\rho^2+\rho^2d\theta^2+dz^2$. Because $\vec{e}_\rho$ and $\vec{e}_\theta$ are independent of z and $\vec{e}_z$ is itself constant, no new non-vanishing Christoffel symbols appear. Now consider a related but different manifold: a cylinder. A cylinder is simply a surface of constant ρ in our three-dimensional Euclidean space. This two-dimensional space is mapped by coordinates (θ,z), with basis vectors $\vec{e}_\theta$ and $\vec{e}_z$. What are the gradients of these basis vectors? They vanish! But, how can that be? From equation (58), $\partial_\theta\vec{e}_\theta = -\rho\vec{e}_\rho$. Have we forgotten about the $\vec{e}_\rho$ direction?

This example illustrates an important lesson. We cannot project tensors into basis vectors that do not exist in our manifold, whether it is a two-dimensional cylinder or a four-dimensional spacetime. A cylinder exists as a two-dimensional mathematical surface whether or not we choose to embed it in a three-dimensional Euclidean space. If it happens that we can embed our manifold into a simpler higher-dimensional space, we do so only as a matter of calculational convenience. If the result of a calculation is a vector normal to our manifold, we must discard this result because this direction does not exist in our manifold. If this conclusion is troubling, consider a cylinder as seen by a two-dimensional ant crawling on its surface. If the ant goes around in circles about the z-axis it is moving in the $\vec{e}_\theta$ direction. The ant would say that its direction is *not* changing as it moves along the circle. We conclude that the Christoffel symbols indeed all vanish for a cylinder described by coordinates (θ, z).

4.4 Gradients of one-forms and tensors

Later we will return to the question of how to evaluate the Christoffel symbols in general. First we investigate the gradient of one-forms and of general tensor fields. Consider a one-form field $\tilde{A}_{\mathbf{X}} = A_{\mu\mathbf{X}}\tilde{e}^\mu{}_{\mathbf{X}}$. Its gradient in a coordinate basis is

$$\tilde{\nabla}\tilde{A} = \tilde{\nabla}(A_\nu\tilde{e}^\nu) = \tilde{e}^\mu\partial_\mu(A_\nu\tilde{e}^\nu) = (\partial_\mu A_\nu)\,\tilde{e}^\mu\tilde{e}^\nu + A_\nu\tilde{e}^\mu\,(\partial_\mu\tilde{e}^\nu) \equiv (\nabla_\mu A_\nu)\,\tilde{e}^\mu\tilde{e}^\nu \ . \tag{60}$$

Again we have defined the covariant derivative operator to give the components of the gradient, this time of the one-form. We cannot assume that ∇_μ has the same form here as in equation (57). However, we can proceed as we did before to determine its relation, if any, to the Christoffel symbols. We note that the partial derivative of a one-form in equation (60) must be a linear combination of one-forms:

$$\partial_\mu\tilde{e}^\nu{}_{\mathbf{X}} \equiv \Pi^\nu{}_{\mu\lambda\,\mathbf{X}}\,\tilde{e}^\lambda{}_{\mathbf{X}} \ , \tag{61}$$

for some set of coefficients $\Pi^\nu{}_{\mu\lambda}$ analogous to the Christoffel symbols. In fact, these coefficients are simply related to the Christoffel symbols, as we may see by differentiating the scalar product of dual basis one-forms and vectors:

$$0 = \partial_\mu\langle\tilde{e}^\nu, \vec{e}_\lambda\rangle = \Pi^\nu{}_{\mu\kappa}\langle\tilde{e}^\kappa, \vec{e}_\lambda\rangle + \Gamma^\kappa{}_{\mu\lambda}\langle\tilde{e}^\nu, \vec{e}_\kappa\rangle = \Pi^\nu{}_{\mu\lambda} + \Gamma^\nu{}_{\mu\lambda} \ . \tag{62}$$

We have used equation (13) plus the linearity of the scalar product. The result is $\Pi^\nu{}_{\mu\lambda} = -\Gamma^\nu{}_{\mu\lambda}$, so that equation (61) becomes, simply,

$$\partial_\mu\tilde{e}^\nu{}_{\mathbf{X}} = -\Gamma^\nu{}_{\mu\lambda\,\mathbf{X}}\,\tilde{e}^\lambda{}_{\mathbf{X}} \ . \tag{63}$$

Consequently, the components of the gradient of a one-form $\tilde{A}$, also known as the covariant derivative of A_ν, are

$$\nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma^\lambda{}_{\mu\nu}A_\lambda \ . \tag{64}$$

This expression is similar to equation (57) for the covariant derivative of a vector except for the sign change and the exchange of the indices ν and λ on the Christoffel symbol (obviously necessary for consistency with tensor index notation). Although we still don't know the values of the Christoffel symbols in general, at least we have introduced no more unknown quantities.

We leave it as an exercise for the reader to show that extending the covariant derivative to higher-rank tensors is straightforward. First, the partial derivative of the components is taken. Then, one term with a Christoffel symbol is added for every index on the tensor component, with a positive sign for contravariant indices and a minus sign for covariant indices. That is, for a (m, n) tensor, there are m positive terms and n negative terms. The placement of labels on the Christoffel symbols is a straightforward extension of equations (57) and (64). We illustrate this with the gradients of the $(0, 2)$ metric tensor, the $(1, 1)$ identity tensor and the $(2, 0)$ inverse metric tensor:

$$\tilde{\nabla}\mathsf{g} = (\nabla_\lambda g_{\mu\nu})\,\tilde{e}^\lambda \otimes \tilde{e}^\mu \otimes \tilde{e}^\nu \ , \quad \nabla_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - \Gamma^\kappa{}_{\lambda\mu} g_{\kappa\nu} - \Gamma^\kappa{}_{\lambda\nu} g_{\mu\kappa} \ , \tag{65}$$

$$\tilde{\nabla}\mathsf{I} = (\nabla_\lambda \delta^\mu{}_\nu)\,\tilde{e}^\lambda \otimes \vec{e}_\mu \otimes \tilde{e}^\nu \ , \quad \nabla_\lambda \delta^\mu{}_\nu = \partial_\lambda \delta^\mu{}_\nu + \Gamma^\mu{}_{\lambda\kappa} \delta^\kappa{}_\nu - \Gamma^\kappa{}_{\lambda\nu} \delta^\mu{}_\kappa \ , \tag{66}$$

and

$$\tilde{\nabla}\mathsf{g}^{-1} = (\nabla_\lambda g^{\mu\nu})\,\tilde{e}^\lambda \otimes \vec{e}_\mu \otimes \vec{e}_\nu \ , \quad \nabla_\lambda g^{\mu\nu} = \partial_\lambda g^{\mu\nu} + \Gamma^\mu{}_{\lambda\kappa} g^{\kappa\nu} + \Gamma^\nu{}_{\lambda\kappa} g^{\mu\kappa} \ . \tag{67}$$

Examination of equation (66) shows that the gradient of the identity tensor vanishes identically. While this result is not surprising, it does have important implications. Recall from Section 2 the isomorphism between g, I and g^{-1} (eq. 47). As a result of this isomorphism, we would expect that all three tensors have vanishing gradient. Is this really so?

For a smooth (differentiable) manifold the gradient of the metric tensor (and the inverse metric tensor) indeed vanishes. The proof is sketched as follows. At a given point $\mathbf{x}$ in a smooth manifold, we may construct a locally flat orthonormal (Cartesian) coordinate system. We define a *locally flat coordinate system* to be one whose coordinate basis vectors satisfy the following conditions in a finite neighborhood around $\mathbf{X}$: $\vec{e}_{\mu\mathbf{X}} \cdot \vec{e}_{\nu\mathbf{X}} = 0$ for $\mu \neq \nu$ and $\vec{e}_{\mu\mathbf{X}} \cdot \vec{e}_{\mu\mathbf{X}} = \pm 1$ (with no implied summation).

The existence of a locally flat coordinate system may be taken as the definition of a smooth manifold. For example, on a two-sphere we may erect a Cartesian coordinate system $x^{\bar{\mu}}$, with orthonormal basis vectors $\vec{e}_{\bar{\mu}}$, applying over a small region around $\mathbf{x}$. (We use a bar to indicate the locally flat coordinates.) While these coordinates cannot, in general, be extended over the whole manifold, they are satisfactory for measuring distances in the neighborhood of $\mathbf{x}$ using equation (41) with $g_{\bar{\mu}\bar{\nu}} = \eta_{\bar{\mu}\bar{\nu}} = g^{\bar{\mu}\bar{\nu}}$, where $\eta_{\bar{\mu}\bar{\nu}}$ is the metric of a flat space or spacetime with orthonormal coordinates (the Kronecker delta or the Minkowski metric as the case may be). The key point is that this statement is true not only at $\mathbf{x}$ but also in a small neighborhood around it. (This argument relies on the absence of curvature singularities in the manifold and would fail, for example, if it were applied at the tip of a cone.) Consequently, the metric must have vanishing first derivative at $\mathbf{x}$ in the locally flat coordinates: $\partial_{\bar{\lambda}} g_{\bar{\mu}\bar{\nu}} = 0$. The gradient of the metric (and the inverse metric) vanishes in the locally flat coordinate basis. But, the gradient of the metric is a tensor and tensor equations are true in any basis. Therefore, for any smooth manifold,

$$\tilde{\nabla}\mathsf{g} = \tilde{\nabla}\mathsf{g}^{-1} = 0 \ . \tag{68}$$

4.5 Evaluating the Christoffel symbols

We can extend the argument made above to prove the symmetry of the Christoffel symbols: $\Gamma^{\lambda}{}_{\mu\nu} = \Gamma^{\lambda}{}_{\nu\mu}$ for any coordinate basis. At point $\mathbf{x}$, the basis vectors corresponding to our locally flat coordinate system have vanishing derivatives: $\partial_{\bar{\mu}} \vec{e}_{\bar{\nu}} = 0$. From equation (56), this implies that *the Christoffel symbols vanish at a point in a locally flat coordinate basis.* Now let us transform to any other set of coordinates x^{μ}. The Jacobian of this transformation is $\Lambda^{\kappa}{}_{\bar{\mu}} = \partial x^{\kappa}/\partial x^{\bar{\mu}}$. Our basis vectors transform (eq. 28) according to $\vec{e}_{\bar{\mu}} = \Lambda^{\kappa}{}_{\bar{\mu}} \vec{e}_{\kappa}$. We now evaluate $\partial_{\bar{\mu}} \vec{e}_{\bar{\nu}} = 0$ using the new basis vectors, being careful to use equation (56) for their partial derivatives (which do *not* vanish in non-flat coordinates):

$$0 = \partial_{\bar{\mu}} \vec{e}_{\bar{\nu}} = \frac{\partial^2 x^{\kappa}}{\partial x^{\bar{\mu}} \partial x^{\bar{\nu}}}\, \vec{e}_{\kappa} + \frac{\partial x^{\kappa}}{\partial x^{\bar{\mu}}} \frac{\partial x^{\lambda}}{\partial x^{\bar{\nu}}}\, \Gamma^{\sigma}{}_{\kappa\lambda} \vec{e}_{\sigma} = 0 \ . \tag{69}$$

Exchanging $\bar{\mu}$ and $\bar{\nu}$ we see that

$$\Gamma^{\sigma}{}_{\kappa\lambda} = \Gamma^{\sigma}{}_{\lambda\kappa} \quad \text{in a coordinate basis} \ , \tag{70}$$

implying that our connection is torsion-free (Wald 1984).

We can now use equations (65), (68) and (70) to evaluate the Christoffel symbols in terms of partial derivatives of the metric coefficients in any coordinate basis. We write $\nabla_{\lambda} g_{\mu\nu} = 0$ and permute the indices twice, combining the results with one minus sign and using the inverse metric at the end. The result is

$$\Gamma^{\lambda}{}_{\mu\nu} = \frac{1}{2} g^{\lambda\kappa} \left(\partial_{\mu} g_{\nu\kappa} + \partial_{\nu} g_{\mu\kappa} - \partial_{\kappa} g_{\mu\nu} \right) \quad \text{in a coordinate basis} \ . \tag{71}$$

Although the Christoffel symbols vanish at a point in a locally flat coordinate basis, they do not vanish in general. This confirms that the Christoffel symbols are not tensor components: If the components of a tensor vanish in one basis they must vanish in all bases.

We can now summarize the conditions defining a locally flat coordinate system $x^{\bar{\mu}}{}_{\mathbf{X}}$ about point $\mathbf{x_0}$: $g_{\bar{\mu}\bar{\nu}\,\mathbf{x_0}} = \eta_{\bar{\mu}\bar{\nu}}$ and $\Gamma^{\bar{\mu}}{}_{\bar{\kappa}\bar{\lambda}\,\mathbf{x_0}} = 0$ or, equivalently, $\partial_{\bar{\mu}} g_{\bar{\kappa}\bar{\lambda}\,\mathbf{x_0}} = 0$.

4.6 Transformation to locally flat coordinates

We have derived an expression for the Christoffel symbols beginning from a locally flat coordinate system. The problem may be turned around to determine a locally flat coordinate system at point $\mathbf{x_0}$, given the metric and Christoffel symbols in any coordinate system. The coordinate transformation is found by expanding the components $g_{\mu\nu\,\mathbf{X}}$ of the metric in the non-flat coordinates x^μ in a Taylor series about $\mathbf{x_0}$ and relating them to the metric components $\eta_{\bar\mu\bar\nu}$ in the locally flat coordinates $x^{\bar\mu}$ using equation (34):

$$g_{\mu\nu\,\mathbf{X}} = g_{\mu\nu\,\mathbf{x_0}} + (x^\lambda - x_0^\lambda)\,\partial_\lambda g_{\mu\nu\,\mathbf{x_0}} + O(x-x_0)^2 = \eta_{\bar\mu\bar\nu}\frac{\partial x^{\bar\mu}}{\partial x^\mu}\frac{\partial x^{\bar\nu}}{\partial x^\nu} + O(x-x_0)^2 \ . \tag{72}$$

Note that the partial derivatives of $\eta_{\bar\mu\bar\nu}$ vanish as do those of any correction terms to the metric in the locally flat coordinates at $x^{\bar\mu} = x_0^{\bar\mu}$. Equation (72) imposes the two conditions required for a locally flat coordinate system: $g_{\bar\mu\bar\nu\,\mathbf{x_0}} = \eta_{\bar\mu\bar\nu}$ and $\partial_{\bar\mu} g_{\bar\kappa\bar\lambda\,\mathbf{x_0}} = 0$. However, the *second* partial derivatives of the metric do not necessarily vanish, implying that we cannot necessarily make the derivatives of the Christoffel symbols vanish at $\mathbf{x_0}$. Quadratic corrections to the flat metric are a manifestation of curvature. In fact, we will see that all the information about the curvature and global geometry of our manifold is contained in the first and second derivatives of the metric. But first we must see whether general coordinates x^μ can be transformed so that the zeroth and first derivatives of the metric at $\mathbf{x_0}$ match the conditions implied by equation (72).

We expand the desired locally flat coordinates $x^{\bar\mu}$ in terms of the general coordinates x^μ in a Taylor series about the point $\mathbf{x_0}$:

$$x^{\bar\mu} = x_0^{\bar\mu} + A^{\bar\mu}{}_\kappa(x^\kappa - x_0^\kappa) + B^{\bar\mu}{}_{\kappa\lambda}(x^\kappa - x_0^\kappa)(x^\lambda - x_0^\lambda) + O(x-x_0)^3 \ , \tag{73}$$

where $x_0^{\bar\mu}$, $A^{\bar\mu}{}_\kappa$ and $B^{\bar\mu}{}_{\kappa\lambda}$ are all constants. We leave it as an exercise for the reader to show, by substituting equations (73) into equations (72), that $A^{\bar\mu}{}_\kappa$ and $B^{\bar\mu}{}_{\kappa\lambda}$ must satisfy the following constraints:

$$g_{\kappa\lambda\,\mathbf{x_0}} = \eta_{\bar\mu\bar\nu} A^{\bar\mu}{}_\kappa A^{\bar\nu}{}_\lambda \ , \quad B^{\bar\mu}{}_{\kappa\lambda} = \frac{1}{2} A^{\bar\mu}{}_\mu \Gamma^\mu{}_{\kappa\lambda\,\mathbf{x_0}} \ . \tag{74}$$

If these constraints are satisfied then we have found a transformation to a locally flat coordinate system. It is possible to satisfy these constraints provided that the metric and the Christoffel symbols are finite at $\mathbf{x_0}$. This proves the consistency of the assumption underlying equation (68), at least away from singularities. (One should not expect to find a locally flat coordinate system centered on a black hole.)

From equation (74), we see that for a given matrix $A^{\bar\mu}{}_\kappa$, $B^{\bar\mu}{}_{\kappa\lambda}$ is completely fixed by the Christoffel symbols in our nonflat coordinates. So, the Christoffel symbols determine the quadratic corrections to the coordinates relative to a locally flat coordinate system. As for the $A^{\bar\mu}{}_\kappa$ matrix giving the linear transformation to flat coordinates, it has 16 independent coefficients in a four-dimensional spacetime. The metric tensor has only 10 independent coefficients (because it is symmetric). From equation (74), we see that we are left with 6 degrees of freedom for any transformation to locally flat spacetime coordinates. Could these 6 have any special significance? Yes! Given any locally flat coordinates in spacetime, we may rotate the spatial coordinates by any amount (labeled by one angle) about any direction (labeled by two angles), accounting for three degrees of freedom. The other three degrees of freedom correspond to a rotation of one of the space coordinates with the time coordinate, i.e., a Lorentz boost! This is exactly the freedom we would expect in defining an inertial frame in *special* relativity. Indeed, in a locally inertial frame general relativity reduces to special relativity by the Equivalence Principle.

4.7 Volume integration

Before proceeding further with differentiation let us consider some aspects of integration over a curved spacetime manifold. Suppose that we want to integrate a scalar field over a four-dimensional spacetime volume. We then need the *proper* four-volume element, i.e., the physical volume element that is invariant (a scalar) under coordinate transformations. We would like to express the proper volume element in terms of coordinate differentials dx^μ. In locally flat coordinates the proper volume element is $d^4x \equiv dx^{\bar{0}}dx^{\bar{1}}dx^{\bar{2}}dx^{\bar{3}}$. However, d^4x is clearly not a scalar under general coordinate transformations. Fortunately, it is easy enough to obtain the correct scalar by looking for a scalar proper volume element that reduces to d^4x for locally flat coordinates. We recall from elementary calculus that if we transform coordinates, the volume element is multiplied by the determinant of the Jacobian matrix for the transformation. So, transforming from locally flat coordinates $x^{\bar{\mu}}$ to general coordinates x^μ, the proper volume becomes

$$d\omega \equiv \left|\det\left[\frac{\partial x^{\bar{\mu}}}{\partial x^\mu}\right]\right| \, d^4x \ . \tag{75}$$

Now, using the first of equations (74), we determine the Jacobian determinant from the metric. Using matrix notation, the metric transformation relation is $g = A^{\mathrm{T}}\eta A$ where A is the Jacobian matrix appearing in equation (75) and A^{T} is its transpose. We now use the theorem of determinants: the determinant of a product of matrices equals the product of determinants. The determinant of $\eta_{\bar{\mu}\bar{\nu}} = \mathbf{diag}(-1,+1,+1,+1)$ is -1 for a four-dimensional spacetime. (If we are considering a spatial manifold with no time, the determinant is $+1$.) Defining $g = \det[g_{\mu\nu}]$ to be the determinant of the metric, we find that $\det A = \sqrt{-g}$ and so the proper volume element is

$$d\omega = \sqrt{-g}\, d^4x \ , \tag{76}$$

This result is easy to understand in a coordinate basis (which we are assuming). From equation (41), $g_{\mu\mu}(dx^\mu)^2$ for a fixed μ (no summation) is the squared distance corresponding to coordinate differential dx^μ. If the coordinates are locally orthogonal the product over μ then gives the square of the volume element. The product of diagonal elements of the metric is not invariant under a rotation of coordinates but the determinant is and it equals the product for a diagonal matrix. For our example of polar coordinates in the Euclidean plane equation (76) recovers the well-known result $d\omega = \rho\, d\rho\, d\theta$.

4.8 Surface integration

Next, suppose that we want to integrate not over a four-dimensional spacetime volume but rather over a three-dimensional volume at fixed time. We can generalize this by asking for the three-dimensional proper "hypersurface" element normal to any given vector ($\vec{e}_t$ in the case of a spatial volume element at fixed t). What does it mean for a vector to be normal to a hypersurface in a curved manifold? The meaning is practically the same as in Euclidean geometry except that the normal is a one-form rather $\tilde{n}$ than a vector. The hypersurface is spanned by a set of vectors $\vec{t}$, called tangent vectors, that are orthogonal at $\mathbf{x}$ to the normal $\tilde{n}$: $\langle \tilde{n}, \vec{t}\,\rangle = 0$. The corresponding normal *vector* $\vec{n}$ (such that $\vec{n}\cdot\vec{t} = 0$) is obtained using the inverse metric (eq. 10).

To find the proper hypersurface element let us first consider the simpler case of a two-dimensional surface in three-dimensional Euclidean space. The surface area element is written $d\vec{S} = \hat{n}\, dA$ where $\hat{n}$ is a unit vector and $dA = d^2x$ (the product of two orthonormal coordinate differentials) is the area normal to $\hat{n}$. We may expect a similar expression in a general spacetime with $\hat{n}$ replaced by the normal one-form $\tilde{n}$ and d^2x replaced by d^3x. However, the determinant of the metric must appear also. To see how, let us choose our coordinates so that $n^0 = (\pm\vec{n}\cdot\vec{n}\,)^{1/2}$ (with a minus sign if $\vec{n}$ is timelike) and $n^i = 0$ for $i = 1, 2, 3$. (It is a simple exercise to show that this is always possible provided that $\vec{n}$ is not a null vector.) In these coordinates, $g_{00} = \pm 1$ and $g_{0i} = 0$ for $i = 1, 2, 3$. We will assume nothing about g_{ij}, the components of the metric in the hypersurface normal to $\tilde{n}$. We know therefore from analogy with equation (76) that to get the proper hypersurface area we must multiply the coordinate hypersurface area by $|g^{(3)}|^{1/2}$ where $g^{(3)}$ is the determinant of g_{ij}. But, our four-by-four metric is block diagonal ($g_{0i} = g_{i0} = 0$) so that the full determinant is $g = g_{00}\, g^{(3)} = \pm g^{(3)}$. (If $\vec{n}$ is timelike, $g^{(3)} > 0$ and the minus sign is taken; if $\vec{n}$ is spacelike, $g^{(3)} < 0$ and the plus sign is taken. In either case, in a four-dimensional spacetime, $g < 0$.) We can therefore write the proper hypersurface element as

$$d\tilde{\sigma} = \tilde{n}\, |g^{(3)}|^{1/2}\, d^3x = \tilde{n}\, \sqrt{-g}\, d^3x \ . \tag{77}$$

This is the desired result.

4.9 Gauss's law

The normal hypersurface element arises naturally in Gauss's law just as it does in the standard vector calculus. To derive Gauss's law we first obtain a general expression for the divergence of a vector field. From equation (57),

$$\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\lambda{}_{\lambda\mu} V^\mu \ . \tag{78}$$

From equation (71) we get $\Gamma^\lambda{}_{\lambda\mu} = (1/2)\, g^{\lambda\kappa}\partial_\mu g_{\lambda\kappa}$. Starting from the definition of the determinant in terms of minors and cofactors (see any linear algebra text), one can show that the logarithmic derivative of the determinant of g is $\partial_\mu \ln g = g^{\lambda\kappa}\partial_\mu g_{\lambda\kappa}$. Combining these results with equation (78) gives us an alternative form for the divergence of a vector field:

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{-g}}\, \partial_\mu(\sqrt{-g}\, V^\mu) \ . \tag{79}$$

The reader may easily verify that this gives the expected results in polar coordinates, taking care to recall that the standard formulae assume vector components in an orthonormal basis rather than a coordinate basis. From equation (79) we obtain the general form of Gauss's law by integrating over proper volume (eq. 76):

$$\int \nabla_\mu V^\mu \, d\omega = \int \partial_\mu(\sqrt{-g}\, V^\mu)\, d^4x = \oint V^\mu n_\mu \sqrt{-g}\, d^3x \ . \tag{80}$$

The four-dimensional integral is a total derivative, allowing it to be reduced to a three-dimensional integral over the closed hypersurface with normal one-form $\tilde{n} = n_\mu \tilde{e}^\mu$ bounding the four-volume region of integration (e.g., $\tilde{n} = \tilde{e}^t$ if we compute the difference of three-volume integrals at two times). Although it is difficult to visualize a four-volume or a closed three-dimensional hypersurface, the meaning is the same as for the standard Gauss's law in space three dimensions except that one more dimension has been added. It is important to note that equation (80) is valid for any coordinate system.

5 Parallel transport and geodesics

5.1 Differentiation along a curve

As a prelude to parallel transport we consider another form of differentiation: differentiation along a curve. A curve is a parametrized path through spacetime: $\mathbf{x}(\tau)$, where τ is a parameter that varies smoothly and monotonically along the path. The curve has a tangent vector $\vec{V} \equiv d\vec{x}/d\tau = (dx^\mu/d\tau)\, \vec{e}_\mu$. If we wish, we could make $\vec{V}$ a unit vector (provided $\vec{V}$ is non-null) by setting $d\tau = (\pm d\vec{x} \cdot d\vec{x}\,)^{1/2}$ to measure path length along the curve (with a minus sign if $d\vec{x}$ is timelike). However, we will impose no such restriction in general. Now, suppose that we have a scalar field $f_{\mathbf{x}}$ defined along the curve (if not all of spacetime). We define the derivative along the curve by a simple extension of equations (36) and (38):

$$\frac{df}{d\tau} \equiv \nabla_V f \equiv \langle \tilde{\nabla} f, \vec{V} \rangle = V^\mu \partial_\mu f \ , \quad \vec{V} = \frac{d\vec{x}}{d\tau}. \tag{81}$$

We have introduced the symbol ∇_V for the covariant derivative along $\vec{V}$, the tangent vector to the curve $\mathbf{x}(\tau)$. This is a natural generalization of ∇_μ, the covariant derivative along the basis vector $\vec{e}_\mu$.

For the derivative of a scalar field, ∇_V involves just the partial derivatives ∂_μ. Suppose, however, that we differentiate a vector field $\vec{A}_{\mathbf{x}}$ along the curve. Now the components of the gradient $\nabla_\mu A^\nu$ are not simply the partial derivatives. As we saw earlier (eq. 57), the covariant derivative of a vector field differs from the partial derivatives. The same is true when we project the gradient onto the tangent vector $\vec{V}$ along a curve:

$$\frac{d\vec{A}}{d\tau} \equiv \frac{DA^\mu}{D\tau}\,\vec{e}_\mu \equiv \nabla_V \vec{A} \equiv \langle \tilde{\nabla}\vec{A}, \vec{V} \rangle = V^\kappa (\nabla_\kappa A^\mu)\,\vec{e}_\mu = \left(\frac{dA^\mu}{d\tau} + \Gamma^\mu{}_{\kappa\lambda} V^\kappa A^\lambda \right) \vec{e}_\mu \ . \tag{82}$$

We retain the symbol ∇_V to indicate the covariant derivative along $\vec{V}$ but we have introduced the new notation $D/D\tau = V^\mu \nabla_\mu \neq d/d\tau = V^\mu \partial_\mu$.

5.2 Parallel transport

The derivative of a vector along a curve leads us to an important concept called parallel transport. Suppose that we have a curve $\mathbf{x}(\tau)$ with tangent $\vec{V}$ and a vector $\vec{A}(0)$ defined at one point on the curve (call it $\tau = 0$). We define a procedure called parallel transport by defining a vector $\vec{A}(\tau)$ along each point of the curve in such a way that $DA^\mu/D\tau = 0$:

$$\nabla_V \vec{A} = 0 \quad \Leftrightarrow \quad \text{parallel transport of } \vec{A} \text{ along } \vec{V}. \tag{83}$$

Over a small distance interval this procedure is equivalent to transporting the vector $\vec{A}$ along the curve in such a way that the vector remains parallel to itself with constant length: $\vec{A}(\tau+\Delta\tau) = \vec{A}(\tau) + O(\Delta\tau)^2$. In a locally flat coordinate system, with Christoffel symbols vanishing at $\mathbf{x}(\tau)$, the components of the vector do not change as the vector is transported along the curve. If the space were globally flat, the components would not change at all no matter how the vector is transported. This is not the case in a curved space or in a flat space with curvilinear coordinates.

5.3 Geodesics

Parallel transport can be used to define a special class of curves called *geodesics*. A geodesic curve is one that parallel-transports its own tangent vector $\vec{V} = d\vec{x}/d\tau$, i.e., a curve that satisfies $\nabla_V \vec{V} = 0$. In other words, not only is $\vec{V}$ kept parallel to itself (with constant magnitude) along the curve, the curve continues to point in the same direction all along the path. A geodesic is the natural extension of the definition of a "straight line" to a curved manifold. Using equations (82) and (83), we get a second-order differential equation for the coordinates of a geodesic curve:

$$\frac{DV^\mu}{D\tau} = \frac{dV^\mu}{d\tau} + \Gamma^\mu{}_{\kappa\lambda} V^\kappa V^\lambda = 0 \quad \text{for a geodesic} \ , \quad V^\mu \equiv \frac{dx^\mu}{d\tau} \ . \tag{84}$$

Indeed, in locally flat coordinates (such that the Christoffel symbols vanish at a point), this is the equation of a straight line. However, in a curved space the Christoffel symbols cannot be made to vanish everywhere. A well-known example of a geodesic in a curved space is a great circle on a sphere.

There are several technical points worth noting about geodesic curves. The first is that $\vec{V} \cdot \vec{V} = \mathsf{g}(\vec{V}, \vec{V})$ is constant along a geodesic because $d\vec{V}/d\tau = 0$ (eq. 84) and $\nabla_V \mathsf{g} = 0$ (eq. 68). Therefore, a geodesic may be classified by its tangent vector as being either timelike ($\vec{V} \cdot \vec{V} < 0$), spacelike ($\vec{V} \cdot \vec{V} > 0$) or null ($\vec{V} \cdot \vec{V} = 0$). The second point is that a nonlinear transformation of the parameter τ will invalidate equation (84). In other words, if $x^\mu(\tau)$ solves equation (84), $y^\mu(\tau) \equiv x^\mu(\xi(\tau))$ will not solve it unless $\xi = a\tau + b$ for some constants a and b. Only a special class of parameters, called *affine parameters*, can parametrize geodesic curves.

The affine parameter has a special interpretation for a non-null geodesic. We deduce this relation from the constancy along the geodesic of $\vec{V} \cdot \vec{V} = (d\vec{x} \cdot d\vec{x})/(d\tau^2) \equiv a$, implying $ds = a d\tau$ and therefore $s = a\tau + b$ where s is the path length (eq. 40). For a non-null geodesic ($\vec{V} \cdot \vec{V} \neq 0$), all affine parameters are linear functions of path length. The linear scaling of path length amounts simply to the freedom to change units of length and to choose any point as $\tau = 0$. Note that originally we imposed no constraints on the parameter but that, for a non-null geodesic, the parameter naturally turns out to measure path length. In fact, we don't need to worry about the parameter in any case because it will take care of itself (even for a null geodesic) when equation (84) is integrated. One can always reparametrize the solution later or eliminate the parameter altogether by replacing it with one of the coordinates along the geodesic.

Another interesting point is that the total path length is stationary for a geodesic:

$$\delta \int_A^B ds = \delta \int_A^B \left(\pm g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right)^{1/2} d\tau = 0 \ . \tag{85}$$

The δ refers to a variation of the integral arising from a variation of the curve, $x^\mu(\tau) \to x^\mu(\tau) + \delta x^\mu(\tau)$, with fixed endpoints. The metric components are considered here to be functions of the coordinates. We leave it as an exercise for the reader to show that equation (85) leads to equation (84). Equation (85) is a curved space generalization of the statement that a straight line is the shortest path between two points in flat space.

5.4 Integrals of motion and Killing vectors

Equation (84) is a set of four second-order nonlinear ordinary differential equations for the coordinates of a geodesic curve. One may ask whether the order of this system can be reduced by finding integrals of the motion. An integral, also called a conserved quantity, is a function of x^μ and $V^\mu = dx^\mu/d\tau$ that is constant along any geodesic. At least one integral always exists: $\vec{V}\cdot\vec{V} = g_{\mu\nu}V^\mu V^\nu$. Are there others? Sometimes. One may show that equation (84) may be rewritten as an equation of motion for $V^\mu \equiv g_{\mu\nu}V^\nu$, yielding

$$\frac{DV^\mu}{D\tau} = \frac{1}{2}(\partial_\mu g_{\kappa\lambda})V^\kappa V^\lambda \ . \tag{86}$$

Consequently, if the metric components are independent of some coordinate x^μ, the corresponding component of the tangent one-form is constant along the geodesic. This result is very useful in reducing the amount of integration needed to construct geodesics for metrics with high symmetry. However, the condition $\partial_\mu g_{\kappa\lambda} = 0$ is coordinate-dependent. There is an equivalnet coordinate-free test for integrals, based on the existence of special vector fields $\vec{\xi}$ call *Killing vectors*. Killing vectors are, by definition, solutions of the differential equation

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0 \ . \tag{87}$$

(The Killing *vector* components are, of course, $\xi^\mu = g^{\mu\nu}\xi_\nu$.) The Killing equation (87) usually has no solutions, but for highly symmetric spacetime manifolds there may be one or more solutions. It is a nice exercise to show that each Killing vector leads to the integral of motion

$$\langle \tilde{V}, \vec{\xi}\rangle = \xi^\mu V_\mu = \text{constant along a geodesic} \ . \tag{88}$$

Note that if one of the basis vectors (for some basis) satisfies the Killing equation, then the corresponding component of the tangent one-form is an integral of motion. The test for integrals implied by equation (86) is a special case of the Killing vector test.

6 Curvature

We introduce curvature by considering parallel transport around a general (non-geodesic) closed curve. In a nonflat space this procedure can lead to a different vector at the end of the curve than the one at the beginning! Consider, for example, a sphere. Suppose we have a vector pointing east on the equator at longitude 0°. We parallel transport the vector eastward on the equator by 180°. At each point on the equator the vector points east. Now the vector is parallel transported along a line of constant longitude over the pole and back to the starting point. At each point on this second part of the curve, the vector points at right angles to the curve, and its direction never changes. Yet, at the end of the curve, at the same point where the curve started, the vector points west!

This result also implies that parallel transport along two different curves may lead to different final vectors. If a vector on the equator is parallel-transported to the opposite side of a sphere along the equator, and, alternatively, over the pole (in fact, both of these curves are geodesics), the resulting vectors will point in opposite directions. This recalls our earlier point that it is not possible to directly compare vectors defined at different points in a curved manifold.

A change in the vector $\vec{A}$ as a result of parallel transport over a closed curve is a dead giveaway of spatial curvature. Suppose that our curve consists of four infinitesimal segments: $d\vec{x}_1$, $d\vec{x}_2$, $-d\vec{x}_1$ and $-d\vec{x}_2$. In a flat space this would be called a parallelogram and the difference $d\vec{A}$ between the final and initial vectors would vanish. In a curved space the change is a linear function of $\vec{A}$, $d\vec{x}_1$ and $d\vec{x}_2$:

$$d\vec{A}(\,\cdot\,) \equiv \mathsf{R}(\,\cdot\,, \vec{A}, d\vec{x}_1, d\vec{x}_2) = \vec{e}_\mu R^\mu{}_{\nu\kappa\lambda}\, A^\nu\, dx^\kappa\, dx^\lambda \ . \tag{89}$$

We have defined a rank $(1,3)$ tensor called the Riemann curvature tensor. (The dots indicate that a one-form is to be inserted; recall that a vector is a function of a one-form.) It is a relatively straightforward exercise to determine the components of the Riemann tensor using equations (82) and (83). We will not go through the algebra here; instead we refer the reader to any differential geometry or general relativity text (e.g., Schutz 1985, section 6.5). The result is

$$R^\mu{}_{\nu\kappa\lambda} = \partial_\kappa \Gamma^\mu{}_{\nu\lambda} - \partial_\lambda \Gamma^\mu{}_{\nu\kappa} + \Gamma^\mu{}_{\alpha\kappa}\Gamma^\alpha{}_{\nu\lambda} - \Gamma^\mu{}_{\alpha\lambda}\Gamma^\alpha{}_{\nu\kappa} \ . \tag{90}$$

Note that some authors (e.g., Weinberg 1972) define the components with opposite sign. Our sign convention follows Misner et al (1973), Wald (1984) and Schutz (1985).

Note that the Riemann tensor involves the first and second partial derivatives of the metric (through the Christoffel symbols). Weinberg (1972) shows that the Riemann tensor is the only tensor that can be constructed from the metric tensor and its first and second partial derivatives and is linear in the second derivatives. Recall that one can always define locally flat coordinates such that $\Gamma^\mu{}_{\nu\lambda} = 0$ at a point. However, one cannot choose coordinates such that $\Gamma^\mu{}_{\nu\lambda} = 0$ everywhere unless the space is globally flat. The Riemann tensor vanishes everywhere if and only if the manifold is globally flat.

An alternative derivation of the Riemann tensor is based on the non-commutativity of the second covariant derivatives of a vector. The covariant derivative $\nabla_\mu A^\nu$ is defined in equation (57); as a rank $(1,1)$ tensor this quantity has a covariant derivative defined as in equation (66). We leave it as an exercise for the reader to show that

$$(\nabla_\kappa \nabla_\lambda - \nabla_\lambda \nabla_\kappa) V^\mu = R^\mu{}_{\nu\kappa\lambda} V^\nu \tag{91}$$

with the Riemann tensor components defined as in equation (90).

If we lower an index on the Riemann tensor components we get the components of a $(0,4)$ tensor:

$$R_{\mu\nu\kappa\lambda} = g_{\mu\alpha} R^{\alpha}{}_{\nu\kappa\lambda} = \frac{1}{2}\left(g_{\mu\lambda,\nu\kappa} - g_{\mu\kappa,\nu\lambda} + g_{\nu\kappa,\mu\lambda} - g_{\nu\lambda,\mu\kappa}\right) + g_{\alpha\beta}\left(\Gamma^{\alpha}{}_{\mu\lambda}\Gamma^{\beta}{}_{\nu\kappa} - \Gamma^{\alpha}{}_{\mu\kappa}\Gamma^{\beta}{}_{\nu\lambda}\right) \ , \tag{92}$$

where we have used commas to denote partial derivatives for notational convenience. In this form it is easy to determine the following symmetry properties of the Riemann tensor:

$$R_{\mu\nu\kappa\lambda} = R_{\kappa\lambda\mu\nu} = -R_{\nu\mu\kappa\lambda} = -R_{\mu\nu\lambda\kappa} \ , \quad R_{\mu\nu\kappa\lambda} + R_{\mu\kappa\lambda\nu} + R_{\mu\lambda\nu\kappa} = 0 \ . \tag{93}$$

It can be shown that these symmetries reduce the number of independent components of the Riemann tensor in four dimensions from 4^4 to 20.

6.1 Bianchi identities, Ricci tensor and Einstein tensor

We must note several more mathematical properties of the Riemann tensor that are needed in general relativity. First, by differentiating the components of the Riemann tensor one can prove the *Bianchi identities*:

$$\nabla_\sigma R^{\mu}{}_{\nu\kappa\lambda} + \nabla_\kappa R^{\mu}{}_{\nu\lambda\sigma} + \nabla_\lambda R^{\mu}{}_{\nu\sigma\kappa} = 0 \ . \tag{94}$$

Note that the gradient symbols denote the covariant derivatives and not the partial derivatives (otherwise we would not have a tensor equation). The Bianchi identities imply the vanishing of the divergence of a certain $(2,0)$ tensor. We first define a symmetric contraction of the Riemann tensor, known as the Ricci tensor:

$$R_{\mu\nu} \equiv R^{\alpha}{}_{\mu\alpha\nu} = R_{\nu\mu} = \partial_\kappa \Gamma^{\kappa}{}_{\mu\nu} - \partial_\mu \Gamma^{\kappa}{}_{\kappa\nu} + \Gamma^{\kappa}{}_{\kappa\lambda}\Gamma^{\lambda}{}_{\mu\nu} - \Gamma^{\kappa}{}_{\mu\lambda}\Gamma^{\lambda}{}_{\kappa\nu} \ . \tag{95}$$

One can show from equations (93) that any other contraction of the Riemann tensor either vanishes or is proportional to the Ricci tensor. The contraction of the Ricci tensor is called the Ricci scalar:

$$R \equiv g^{\mu\nu} R_{\mu\nu} \ . \tag{96}$$

Contracting the Bianchi identities twice and using the antisymmetry of the Riemann tensor one obtains the following relation:

$$\nabla_\nu G^{\mu\nu} = 0 \ , \quad G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = G^{\nu\mu} \ . \tag{97}$$

The symmetric tensor $G^{\mu\nu}$ that we have introduced is called the *Einstein tensor*. Equation (97) is a mathematical identity, not a law of physics. Through the Einstein equations it provides a deep illustration of the connection between mathematical symmetries and physical conservation laws.

References

[1] Bishop, R. L. and Goldberg, S. I. 1980, *Tensor Analysis on Manifolds* (New York: Dover).

[2] arroll, S. M. 1997, gr-qc/9712019.

[3] Frankel, T. 1979, *Gravitational Curvature: An Introduction to Einstein's Theory* (San Francisco: W. H. Freeman).

[4] Lightman, A. P., Press, W. H., Price, R. H. and Teukolsky, S. A. 1975, *Problem Book in Relativity and Gravitation* (Princeton: Princeton University Press).

[5] Lovelock, D. and Rund, H. 1975, *Tensors, Differential Forms, and Variational Principles* (New York: Wiley).

[6] Misner, C. W., Thorne, K. S. and Wheeler, J. A. 1973, *Gravitation* (San Francisco: Freeman).

[7] Schutz, B. F. 1980, *Geometrical Methods of Mathematical Physics* (Cambridge: Cambridge University Press).

[8] Schutz, B. F. 1985, *A First Course in General Relativity* (Cambridge: Cambridge University Press).

[9] Wald, R. M. 1984, *General Relativity* (Chicago: University of Chicago Press).

[10] Weinberg, S. 1972, *Gravitation and Cosmology* (New York: Wiley).

[11] Will, C. M. 1981, *Theory and Experiment in Gravitational Physics* (Cambridge: Cambridge University Press).

[12] Will, C. M. 1986, *Was Einstein Right?* (New York: Basic Books).