Massachusetts Institute of Technology
Department of Physics

Physics 8.962 Spring 2000

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 as a theory of curved spacetime, 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.

In special relativity there exist global inertial frames. This is no longer true in the presence of gravity. However, there are *local* inertial frames in GR, such that within a suitably small spacetime volume around an event (just how small is discussed e.g. in MTW Chapter 1), one may choose coordinates corresponding to a nearly-flat spacetime. Thus, the *local* properties of special relativity carry over to GR. The mathematics of vectors and tensors applies in GR much as it does in SR, with the restriction that vectors and tensors are defined independently at each spacetime event (or within a sufficiently small neighborhood so that the spacetime is sensibly flat).

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, MTW). 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. It expands on MTW Chapters 2, 3, and 8. 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 twodimensional vector space exists at the point of tangency. However, such an embedding is not required to define the tangent space of a manifold (Wald 1984). As long as the space is smooth (as assumed in the formal definition of a manifold), the difference vector $d\vec{x}$ between two 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 later notes we introduce will 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 $\mathbf{g}_{\mathbf{X}}$ called the metric tensor and a related $(2,0)$ tensor field $\mathbf{g}_{\mathbf{X}}^{-1}$ 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 $\vec{W}$, $\mathbf{g}$ returns a scalar, called the dot product:

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

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

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

Although a dot is used in both cases, it should be clear from the context whether $\mathbf{g}$ or $\mathbf{g}^{-1}$ 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. 41) we will see what distinguishes $\mathbf{g}$ from other $(0,2)$ tensors and $\mathbf{g}^{-1}$ 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:

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

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

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

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 $\tilde{W}$:

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

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): $\tilde{e}^\mu(\,\cdot\,) \neq \mathbf{g}(\vec{e}_\mu,\,\cdot\,)$. 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 $\mathbf{g}$ and evaluating the dot product using $\vec{A} \cdot \vec{A} = \mathbf{g}(\vec{A}, \vec{A}) = g_{\mu\nu} A^\mu A^\nu$. The same classification scheme extends to one-forms using $\mathbf{g}^{-1}$: a one-form $\tilde{P}$ is spacelike, timelike or null according to whether $\tilde{P} \cdot \tilde{P} = \mathbf{g}^{-1}(\tilde{P}, \tilde{P}) = g^{\mu\nu} P_\mu P_\nu$ 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 later notes 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.