# A Short Course in General Relativity

## "Course" Notes

## Preface

I taught myself general relativity starting from this text, "A Short Course in General Relativity" by James Foster and J. David Nightingale. These "class notes" started as an elaboration of the book to make it understandable to me, and I then added other topics that I also required. Eventually I integrated methods from other sources, including an MIT course by Scott Hughes to modernize some concepts and notation, especially the geometric underpinnings of General Relativity.

I attempted to fill in ALL of the details that the text glossed over in order to maintain focus on the key threads. I have created my own colored figures that I find much easier to understand than the black and white figures in the book. I have also expanded many sections to include information that I found was necessary for my understanding. For example:

- I have added a chapter 0 with vector space topics, required prerequisites but mostly not found in the book. I have sprinkled other prerequisites, such as expressing multivariate Taylor series in index notation, throughout the book

- I added brief developments of non-relativistic equations for fluid flow, gravitational potential, energy conservation. This includes a couple of equations that Einstein's field equations must reduce to in the limit of Newtonian conditions: Poisson's field equation and Euler's equation of motion for a perfect fluid,

- I have merged appendices A and (part of) C into the main text body at the places where they are needed and naturally fit.

- I have added, modified, expanded, or simplified many topics. For example, the book only gives an example Lorentz transformation consisting of a boost (increase in speed) in the *x*-direction, and then develops the resulting Lorentz transformation matrix. I found it relatively easy to expand that development to the general Lorentz transformation, and the resulting general Lorentz transformation matrix can actually be much simpler to express and easier to understand. I also develop the even simpler homogeneous Lorentz transformation, provide formulas relating the two, and occasionally use the homogeneous transformation to provide yet simpler proofs.

I have maintained the equation numbers used by the book. But.. I have rearranged sections and, thus, many of the book's equations are now a little out-of-order. I have also added numbered equations that I reference. In order to be clear about which equations I have added, I use numbering schemes that differ from the book's. The book numbers equations by chapters, like (4.1) – (4.85) for Chapter 4. I mostly number my equations by section, like (4.1-3) for my third numbered equation of Section 4.1. In addition, I sometimes use simple numbering like (iii) and (a) in localized regions.

# Chapter 0  Vector Spaces

Euclidean spaces, the subject of Section 1-1, are vector spaces. In addition, vector spaces form the backbone of tensor mathematics. This section provides a short review of the vector space topics required for this course. Theorems that are well-known but for which I do not take the time to prove here are labeled as "Facts". Additional vector space basics, needled in Chapter 2, are delayed until Appendix C.2, which I have inserted between Chapters 1 and 2.

Definition  For these notes, **scalars** are either the reals, $\mathbb{R}$, or the complex numbers, $\mathbb{C}$. A **vector space** is a set **V**, whose elements are called **vectors**, in which the operations addition and scalar-multiplication have been defined and satisfy the following natural rules.

1. If v, w, and x are vectors then v + w = w + v and v + (w + x) = (v + w) + x
2. V has a unique vector 0 such that v + 0 = v for every $v \in V$
3. To each $v \in V$ there is a vector -v such that v + (-v) = 0
4. For a scalar $\alpha$, there is a vector $\alpha v \in V$ for each $v \in V$
5. 1v = v, $\alpha(\beta v) = (\alpha\beta)v$ (where 1 is the unit scalar)
6. $\alpha(v + w) = \alpha v + \alpha w$ (0-1)

Definitions

- A set of vectors $\mathcal{V} = \{\boldsymbol{v}_i\}$ is **linearly independent** if $\sum \alpha_i \boldsymbol{v}_i = 0 \Rightarrow \alpha_i = 0$ for all $i$.
- Otherwise we say that $\mathcal{V}$ is **linearly dependent**. (0-2)
- A set of vectors $\mathcal{W} = \{\boldsymbol{w}_i\}$ **spans V** if whenever $\boldsymbol{v} \in \mathbf{V}$, $\exists$ scalars $\alpha_i$ such that $\boldsymbol{v} = \sum \alpha_i \boldsymbol{w}_i$ . (0-3)
- A **basis for a vector space V** is a set $\mathcal{B} = \{\mathbf{e}_i\}$ of linearly independent vectors that spans **V**. The **dimension of V** is the number of elements in $\mathcal{B}$. (0-4)

What this means is that no basis vector can be expressed as a linear combination of the other basis vectors, and that every vector in **V** can be expressed as a linear combination of the basis vectors.

Definition  A **linear transformation** of a vector space V into a vector space W is a mapping $T$: $\mathbf{V} \to \mathbf{W}$ : $T(\alpha\boldsymbol{v} + \beta\boldsymbol{w}) = \alpha\, T(\boldsymbol{v}) + \beta\, T(\boldsymbol{w})$ for all vectors $\boldsymbol{v}$, $\boldsymbol{w} \in \mathbf{V}$ and all scalars $\alpha$ and $\beta$. (0-5)

We describe this by saying a linear transformation preserves the vector space structure. This means that addition and scalar multiplication match up:

$$\alpha\boldsymbol{v} + \beta\boldsymbol{w} \;\leftrightarrow\; \alpha\, T(\boldsymbol{v}) + \beta\, T(\boldsymbol{w})$$

Fact 1  A linear transformation defined on the basis vectors of **V** to **W** has a unique extension to a linear transformation from all of **V** to **W**. That is, if $T$ is defined on a basis of **V**, we can consider it to be defined on all of **V**.

The next definition uses Fact 1 to show how a matrix can be considered to be a linear transformation, and vice-versa.

Definition  Let $\boldsymbol{V}$ be any $N$-dimensional vector space with a basis $\{\mathbf{e}_i\}$, $M = (m_{ij})$ an $N \times N$ matrix whose entries are scalars, and $T$ the linear transformation from $\mathbf{V} \to \mathbf{V}$ defined by

$$T(\mathbf{e}_i) = \sum_j m_{ij}\, \mathbf{e}_j \, .$$

We say that **the matrix *M* is associated with *T***. (0-6)

When working with matrices, we represent vectors $\boldsymbol{v} = \begin{pmatrix} \vdots \\ v_j \\ \vdots \end{pmatrix}$ as column vectors. We then consider the matrix expression $M\boldsymbol{v} = \boldsymbol{w}$ to be interchangeable with the *associated* linear transformation expression $T(\boldsymbol{v}) = \boldsymbol{w}$.

Definition  The **transpose** of a column vector $\boldsymbol{v}$ is the row vector $\boldsymbol{v}^T = (\cdots\ v_j\ \cdots)$. (0-7)

Even though general relativity is based upon vector spaces whose scalars are real, the Fundamental Theorem of algebra shows that complex numbers can arise as solutions to polynomial equations that have only real coefficients. For this reason, we must temporarily work with vector spaces over the complex numbers to solve certain problems. The facts and definitions below are for vector spaces over $\mathbb{C}$ but also apply to vector spaces over $\mathbb{R}$.

Fact 2  Let **V** be a vector space over $\mathbb{C}$, $M$ be an $N$x$N$ complex matrix, and $T : \mathbf{V} \to \mathbf{V}$ the associated linear transformation. Then the following are equivalent.

1. $M$ is singular (i.e., dim[ $T(\mathbf{V})$ ] < $N$)
2. $M$ is non-invertible (i.e., there is no inverse matrix)
3. det $M = 0$
4. There is a non-zero vector $\boldsymbol{v}$ in **V** such that $M\boldsymbol{v} = 0$

Definition Let $M = (m_{ij})$ be a complex-valued $N \times N$ matrix.

The **transpose** of $M$ is the matrix $M^T = (m_{ji})$. (0-8)

Matrix $M$ is **symmetric** if $M = M^T$; i.e., if $m_{ij} = m_{ji}$ for all $i$ and $j$. (0-9)

If $z = x + y\,i$, let $\bar{z} = x - y\,i$ represent its complex conjugate.

The **conjugate** of $M$ is the matrix $M^* = (\overline{m}_{ij})$. (0-10)

We will write $M^{T^*}$ as a shortcut for $(M^T)^*$.

Definition Let $M = (m_{ij})$ be a complex-valued $N \times N$ matrix. We say that $\boldsymbol{v}$ is an **eigenvector of *M*** if there is a scalar $\lambda \in \mathbb{C}$ such that $M\boldsymbol{v} = \lambda\boldsymbol{v}$, and $\lambda$ is called an **eigenvalue**.

Theorem 0.1 All eigenvectors and eigenvalues of a real symmetric matrix $M$ are real.

Proof. Suppose $M\boldsymbol{v} = \lambda\boldsymbol{v}$ where $M$ has real entries, $\lambda \in \mathbb{C}$, and $\boldsymbol{v}$ has complex components. Since $M$ is real, $M^* = M$. Since $M$ is symmetric, $M^T = M$. Thus, $M = M^{T^*}$. So,

$$\boldsymbol{v}^{T^*} M = \boldsymbol{v}^{T^*} M^{T^*} = (M\boldsymbol{v})^{T^*} = (\lambda\boldsymbol{v})^{T^*} = \lambda^* (\boldsymbol{v}^{T^*}) = \lambda^* \boldsymbol{v}^{T^*} \qquad (0\text{-}11)$$

$$\boldsymbol{v}^{T^*} M\boldsymbol{v} = \boldsymbol{v}^{T^*}(M\boldsymbol{v}) = \boldsymbol{v}^{T^*}(\lambda\boldsymbol{v}) = \lambda\, \boldsymbol{v}^{T^*}\boldsymbol{v}$$

$$\boldsymbol{v}^{T^*} M\boldsymbol{v} = (\boldsymbol{v}^{T^*} M)\,\boldsymbol{v} \overset{(0-11)}{=} (\lambda^* \boldsymbol{v}^{T^*})\,\boldsymbol{v} = \lambda^* \boldsymbol{v}^{T^*}\boldsymbol{v}$$

$$\Rightarrow \quad \lambda = \lambda^* \quad \Rightarrow \quad \lambda \text{ is real}$$

Because $M$ is a real matrix and its eigenvalues are also real, the eigenvectors must also be real. That is because $M\boldsymbol{v} = \lambda\boldsymbol{v}$ constitutes a system of real linear equations, and solving them only involves addition, subtraction, multiplication, and division. It does not involve taking complex conjugates. ■

Definition The **dot product** of two (column or row) vectors $\boldsymbol{v} = (v_i)$ and $\boldsymbol{w} = (w_i)$ is

$$\boldsymbol{v} \cdot \boldsymbol{w} = \Sigma\, v_i\, w_i.$$

Vectors are **orthogonal** if $\boldsymbol{v} \cdot \boldsymbol{w} = 0$. They are **orthonormal** if in addition they are unit vectors; i.e., $\sum v_i^{\,2} = 1$. The **norm** of a vector is

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}\ .$$

The remainder of this section is devoted to proving the following theorem (needed in proof of Theorem 2.4.2 that every point has a coordinate system where the metric tensor matrix is a diagonal matrix having only +1's and -1's).

Theorem 0.2 A real symmetric matrix $M = (m_{ij})$ has a collection of eigenvectors that constitute an orthonormal basis for **V**.

Proof. Let $I$ be the identity matrix. $p(\lambda) \equiv \det(M - \lambda I)$ is a polynomial in $\lambda$ known as the characteristic polynomial of $M$. By the Fundamental Theorem of Algebra, the polynomial $p(\lambda)$ can be factored,

$$p(\lambda) = \pm (\lambda - \lambda_1)\cdots(\lambda - \lambda_N),$$

where $\lambda_i$ are complex roots (even though the elements of $M$ are real). It is possible that some or all of the $\lambda_i$ are the same, but we have shown that $p(\lambda)$ has at least one characteristic root, $\lambda_1$.

We next show that $M$ has at least one eigenvector; namely, an eigenvector $\boldsymbol{v}$ whose eigenvalue is the characteristic root $\lambda_1$. We seek a non-zero vector $\boldsymbol{v}$ that satisfies $M\boldsymbol{v} = \lambda_1 \boldsymbol{v}$. Denote $\boldsymbol{v}$ as the column vector $(v_j)$.

Let $\{\mathbf{e}_i\}$ be a basis for **V**. By the definition above, the matrix $M = (m_{ij})$ is associated with the linear transformation $T$ defined on the basis vectors of $\boldsymbol{V}$ by

$$T : \mathbf{V} \to \mathbf{V} : T(\mathbf{e}_i) = m_{ij}\, \mathbf{e}_j .$$

Define a matrix $H \equiv M - \lambda_1 I$, where $I$ is the identity matrix (ones on the main diagonal, zeroes elsewhere). So, $\det(H) = \det(M - \lambda_1 I) = p(\lambda_1) = 0$. By Fact 2, there is a non-zero, possibly complex vector $\boldsymbol{w}$ in $\boldsymbol{V}$ such that $H\boldsymbol{w} = 0$. That is,

$$M\boldsymbol{w} - \lambda_1 \boldsymbol{w} = (M - \lambda_1)\,\boldsymbol{w} = H\boldsymbol{w} = 0 \qquad \Rightarrow \qquad M\boldsymbol{w} = \lambda_1 \boldsymbol{w}.$$

Thus, $\boldsymbol{w}$ is an eigenvector of $M$ having the characteristic root $\lambda_1$ as its eigenvalue. By Theorem 0.1, $\boldsymbol{v}$ and $\lambda_1$ are real.

Having found one eigenvector, we now extend it into an orthonormal basis for $\boldsymbol{V}$. We set $\boldsymbol{v}_1 = \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}$, a unit vector having $\lambda_1$ as its eigenvalue. Define the **null space** of $\boldsymbol{v}_1$:

$$N_1 = \{\, \boldsymbol{v} : \boldsymbol{v} \bullet \boldsymbol{v}_1 = 0 \}.$$

It is easy to confirm that $N_1$ satisfies the definition, above, of a vector space. $N_1$ is the subspace of vectors that are orthogonal to $\boldsymbol{v}_1$. Claim dim $(N_1) = N - 1$:

Using the Gram-Schmidt orthogonalization process, $\boldsymbol{v}_1$ can be extended to an orthonormal basis $\{\boldsymbol{v}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_N\}$ of $\boldsymbol{V}$. Thus, $\boldsymbol{e}_i \bullet \boldsymbol{v}_1 = 0$ for all $i > 1$ since $\boldsymbol{e}_i$ and $\boldsymbol{v}_1$ are orthonormal basis vectors. Hence, by the definition of $N_1$, $\{\boldsymbol{e}_2, \cdots, \boldsymbol{e}_N\}$ is contained in $N_1$, and, thus, forms an $(N-1)$ dimensional basis for it. ✓

Claim $MN_1 \subseteq N_1$ where $MN_1 = \{M\boldsymbol{v} : \boldsymbol{v} \in N_1\}$:

Let $\boldsymbol{v} \in N_1$. We need to show that $M\boldsymbol{v} \in N_1$. Since $\boldsymbol{v} \bullet \boldsymbol{v}_1 = \boldsymbol{v}^T\boldsymbol{v}_1$ and $M = M^{T^*}$,

$$M\boldsymbol{v} \bullet \boldsymbol{v}_1 = (M\boldsymbol{v})^{T^*}\boldsymbol{v}_1 = \boldsymbol{v}^{T^*}M\boldsymbol{v}_1 = \boldsymbol{v}^{T^*}\lambda_1\,\boldsymbol{v}_1 = \lambda_1\,\boldsymbol{v}^T\boldsymbol{v}_1 = \lambda_1\,\boldsymbol{v} \bullet \boldsymbol{v}_1 = 0 \qquad ✓$$

Let $T_2$ be the linear transformation generated by restricting $T$ to $N_1$, the $(N-1)$ dimensional null space of $\boldsymbol{v}_1$, and let $M_2$ be the matrix associated with $T_2$. Repeating our logic above, $p(\lambda) \equiv \det(M_2 - \lambda\, I)$ has a real root $\lambda_2$ that is an eigenvalue of $M_2$, and $\lambda_2$ has a corresponding unit eigenvector $\boldsymbol{v}_2$. Because $\boldsymbol{v}_2 \in N_1$, we have $\boldsymbol{v}_2 \bullet \boldsymbol{v}_1 = 0$. $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ forms an orthonormal set. (So, even though $\lambda_2$ might equal $\lambda_1$, we have that $\boldsymbol{v}_2 \neq \boldsymbol{v}_1$.)

We have to do this one more time. Let $T_3$ be the linear transformation generated by restricting $T_2$ to $N_2$, the $(N-2)$ dimensional null space of $\boldsymbol{v}_2$. Let $M_3$ be the matrix associated with $T_3$. As above, we generate a real eigenvalue $\lambda_3$ having a corresponding unit eigenvector $\boldsymbol{v}_3$ that is orthogonal to $\boldsymbol{v}_2$, and since $\boldsymbol{v}_3 \in N_2 \subseteq N_1$, we also have that it is orthogonal to $\boldsymbol{v}_1$. Thus, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ form an orthonormal set.

Continuing this process, we eventually obtain the orthonormal basis $\{\boldsymbol{v}_i\}$ consisting of eigenvectors of $M$ that have corresponding real eigenvalues. ■

# Chapter 1 Vector and Tensor Fields

## 1.0 Introduction

General relativity (GR) is the study of gravity as curvature in spacetime. On a curved surface such as a globe, any small enough region appears flat; i.e., approximately Euclidean. The globe can be covered with overlapping small regions, each with its own coordinate system. As one traverses the globe, it is necessary to move from one local coordinate system to another overlapping system. This chapter sets the groundwork for local coordinate systems and translating from one to another.

There are several global concepts to discuss, upfront.

- **Basis-free**
  - ✓ Vectors and tensors are represented using only the basis coefficients
  - ✓ This is a convenience of GR
- **Coordinate-free**
  - ✓ Expressions and equations written without reference to coordinates
  - ✓ A desirable underpinning for any theory
- **Coordinate-independent**
  - ✓ Expressions and equations that look the same in all coordinate systems
  - ✓ A requirement of GR
- **Principle of general covariance**
  - ✓ GR formulas must be coordinate-independent and reduce locally to the equations of special relativity.

Basis-free. In a vector space, vectors are represented as linear combinations of basis vectors: $\boldsymbol{v} = \Sigma \alpha_i \, \boldsymbol{e}_i$ . The tensors form a vector space and so they, too, are generated from a set of basis vectors. A vector is an object that has magnitude and direction. If we use unit basis vectors, then they represent directions and the coefficients represent the magnitude in each direction. Whether or not the basis vectors are unit vectors, the direction of a vector can be determined from just the magnitudes. Einstein simplified GR notation by primarily using expressions involving just the coefficients, dropping the basis vectors.

Coordinate-free. Geometry is the prime example of a coordinate-free theory. Coordinate-free is a desirable property of any theory. Geometry was solely coordinate-free for two thousand years. It wasn't until Descartes invented Cartesian geometry that a coordinate-dependent geometry appeared, and it has regular geometry as a foundation. Newtonian physics embodies extensive coordinate-based computations but it, too, has a coordinate-free foundation consisting of equations like $F = ma$ . GR was developed entirely within a coordinate framework and lacked a coordinate-free foundation until around year 2000 when a geometry framework was developed (after publication of Foster and Nightingale's book). This framework introduces the objects of geometric algebra, and enables the powerful methods of geometric calculus to be used.

An overview of this process is given in Appendix D (my own–not in book)

Coordinate-independent. For general relativity, the most important of these framework considerations is coordinate-independence. Saying that an expression or equation is coordinate-independent means that it has the same appearance in every coordinate system. For example, let $(x^i)$ and $(x^{i'})$ be two coordinate systems. We will see that the dot-product operation on two vectors is a coordinate-independent expression because it is unchanged after a coordinate transformation:

$$\boldsymbol{v} \bullet \boldsymbol{w} = g_{ij}\, v^i v^j = g_{i' j'}\, v^{i'} v^{j'} = \boldsymbol{v}' \bullet \boldsymbol{w}'.$$

The Dirac delta symbol is is coordinate-independent: $\delta^i_j = \delta^{i'}_{j'}$

Non-GR expressions are not required to be coordinate independent although we point them out when they occur. However, Einstein required coordinate-independence of every general relativity expression. A standard process will be developed for transformation of coordinates (using Jacobian matrices) that preserves coordinate-independence of vectors and tensors. Thereafter, any vector or tensor expression generated by transforming a given vector or tensor expression will automatically be coordinate-independent. However, when proposing new definitions or relationships that are not derived, such as the formula for a geodesic curve or Einstein's field equations, they must be checked to confirm that they satisfy coordinate-independence.

Principle of general covariance. Within small volumes of space and short time intervals, equations must resemble special relativity to a 1st order approximation. This is akin to a small region on Earth appearing flat. Together with coordinate-independence, Einstein called these two properties the **principle of general covariance**.

## 1.1 Euclidean Coordinate Systems (3-Space)

### Definition

• A **Coordinate System** is a system that uses one or more numbers, or coordinates, to uniquely determine the position of the points or other geometric elements on a manifold such as Euclidean space.

The term "unique" means that a coordinate, say ($a$,$b$), represents a unique point, P. It does not mean that the point P has a unique coordinate representation. This clarification becomes important when we consider, for example, whether the polar coordinate system ($r$,$\theta$) is defined at the origin. The origin O can be represented by the many different pairs (0,$\theta$), but as long as each pair of coordinates uniquely determines O, the origin is considered to be a well-defined point in the coordinate system.

### Definitions

• A **Cartesian system** is a coordinate system in which points are represented with coordinates *(x,y,z)* along axes that are mutually orthogonal.
• We use the symbol $\mathbb{R}$ to represent the set of real numbers
• We use the symbol $\mathbb{R}^3$ to represent the set of points P = ($x$,$y$,$z$), where $x,y,z \in \mathbb{R}$
• The **Euclidean metric** is the distance measure for points in $\mathbb{R}^3$ :

$$d(P,Q) \equiv \sqrt{(x-\tilde{x})^2 + (y-\tilde{y})^2 + (z-\tilde{z})^2}\text{, where } Q = (\tilde{x},\ \tilde{y},\ \tilde{z})\ .$$

✓ The metric satisfies d(P,Q) ≥ 0 and d(P,Q) = 0 iff P = Q.

• Rather than the symbol $\mathbb{R}^3$, we use the symbol $\mathbb{E}^3$ to represent **Euclidean 3-space**, the space of points $\mathbb{R}^3$ along with the Euclidean metric.

In chapters 1 and 2 we shall assume a permanent Euclidean 3-space with a fixed *(x,y,z)* Cartesian coordinate system having unit vectors ***i***, ***j***, and ***k*** pointing along the positive $x$, $y$, and $z$ axes, respectively.

We point out in passing that the ***i-j-k*** scheme works only in 3 dimensions, not in 1 dimension like a line, 2 dimensions like a plane or surface of a sphere, or 4 dimensions like spacetime. This is because vector cross products are only defined in 3-space. In particular, we cannot develop a 4-manifold or a spacetime scheme using ***i-j-k-ℓ***.

The modern point-of-view (as of 2024) is that GR is a coordinate-free discipline, having geometric objects that exist independent of any coordinate system but which can be *represented* in coordinate systems. We consider a point to be a geometric object that has a position. We consider a vector to be a geometric object that has magnitude and direction. Both can be assigned coordinates given a coordinate system. When we write P = $(x,y,z)$ , we should be aware that actually mean P $\doteq (x,y,z)$, where $\doteq$ means that the geometric object P is represented by $(x,y,z)$ in a certain coordinate system O. Similarly, when we write a vector object as $\boldsymbol{v} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$, we actually mean $\boldsymbol{v} \doteq x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$.

**Position vectors** $\boldsymbol{r}$ are be expressed in Cartesian coordinates by

$$\boldsymbol{r} = \boldsymbol{r}(x,y,z) = \mathrm{x}\,\boldsymbol{i} + \mathrm{y}\,\boldsymbol{j} + \mathrm{z}\,\boldsymbol{k}\,. \qquad (1.3)$$

When working with curves and curved surfaces in 3-space it is necessary to define overlapping coordinate systems to move from one point to another. Let $(x,y,z)$ be a Cartesian coordinate system and $(u,v,w)$ be an alternate coordinate system, possibly another Cartesian system but usually a non-Cartesian system such as spherical $(r,\theta,\phi)$ or cylindrical $(\rho,\phi,\mathrm{z})$ coordinates. The alternate coordinate system can have a different center and it can also involve a rotation.

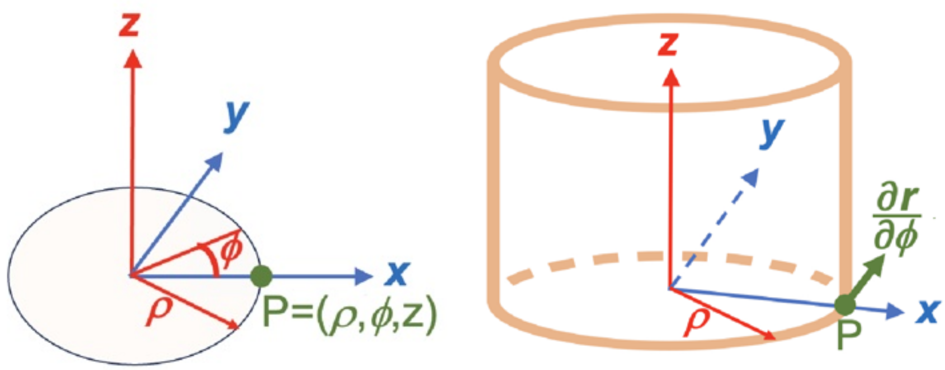
In principle, we can express $(x,y,z)$ in terms of $(u,v,w)$ and $(u,v,w)$ in terms of $(x,y,z)$:

$$x = x(u,v,w), \quad y = y(u,v,w), \quad \text{and} \quad z = z(u,v,w) \qquad (1.1)$$

$$u = u(x,y,z), \quad v = v(x,y,z), \quad \text{and } w = w(x,y,z)\,. \qquad (1.7)$$

Let P = $(u_0,v_0,w_0)$ be a point. Observe that $\{\ (u,v_0,w_0)\ \}$ is a curve in space because only one parameter varies. We will restrict our attention to differentiable curves. Then, $\frac{\partial \boldsymbol{r}}{\partial u}$, evaluated at P, is a vector that is tangent to the curve at P. Similarly, $\frac{\partial \boldsymbol{r}}{\partial v}$ and $\frac{\partial \boldsymbol{r}}{\partial w}$ are vectors that are tangent to space curves $\{\ (u_0,v,w_0)\ \}$ and $\{\ (u_0,v_0,w)\ \}$, respectively, that pass through P.

Example 1.1.1 The figure below on the left shows cylindrical coordinates $(\rho,\phi,z)$ and a point P = $(\rho,\phi,z) = (2,0,0)$ in cylindrical coordinates. By chance, P = (2,0,0) in $xyz$-coordinates, also. When $\phi$ varies, the curve formed is a circle in the $xy$-plane. The right-hand figure shows that the circle lies on the surface of a cylinder of radius 2 about the $z$-axis. The vector $\frac{\partial \boldsymbol{r}}{\partial \phi}$ is in the $xy$-plane. The vector begins at P and points parallel to the positive $y$-axis. In Exercise 1.1.3. We will be able to see that this vector has magnitude 2.




Definition The **natural basis** at a point P = ($u_0$,$v_0$,$w_0$) is defined to be

$$\begin{aligned}
\boldsymbol{e}_u &\equiv \frac{\partial \boldsymbol{r}}{\partial u} \overset{1.3}{=} \frac{\partial x}{\partial u}\boldsymbol{i} + \frac{\partial y}{\partial u}\boldsymbol{j} + \frac{\partial z}{\partial u}\boldsymbol{k} \\
\boldsymbol{e}_v &\equiv \frac{\partial \boldsymbol{r}}{\partial v} = \frac{\partial x}{\partial v}\boldsymbol{i} + \frac{\partial y}{\partial v}\boldsymbol{j} + \frac{\partial z}{\partial v}\boldsymbol{k} \\
\boldsymbol{e}_w &\equiv \frac{\partial \boldsymbol{r}}{\partial w} = \frac{\partial x}{\partial w}\boldsymbol{i} + \frac{\partial y}{\partial w}\boldsymbol{j} + \frac{\partial z}{\partial w}\boldsymbol{k},
\end{aligned} \tag{1.6}$$

where the partials are evaluated at P.

We will show in the Corollary to Theorem 1.2.1 that the natural basis is in fact a basis; i.e., the vectors are linearly independent. The basis {$\boldsymbol{e}_u$,$\boldsymbol{e}_v$,$\boldsymbol{e}_w$} is composed of the tangent vectors to the 1-parameter *space curves* passing through ($u_0$,$v_0$,$w_0$). The basis can possibly contain neither orthogonal nor unit vectors.

Another approach is to use gradients to make a basis composed of normals to the *space surfaces*, like ($u$ ,$v$,$w_0$), passing through ($u_0$,$v_0$,$w_0$). We observe that ($u$ ,$v$,$w_0$) is a surface because 2 parameters vary.

Definition The **dual basis** at a point P is

$$\begin{aligned}
\boldsymbol{e}^u &\equiv \nabla u \overset{1.7}{=} \frac{\partial u}{\partial x}\boldsymbol{i} + \frac{\partial u}{\partial y}\boldsymbol{j} + \frac{\partial u}{\partial z}\boldsymbol{k} \\
\boldsymbol{e}^v &\equiv \nabla v = \frac{\partial v}{\partial x}\boldsymbol{i} + \frac{\partial v}{\partial y}\boldsymbol{j} + \frac{\partial v}{\partial z}\boldsymbol{k} \\
\boldsymbol{e}^w &\equiv \nabla w = \frac{\partial w}{\partial x}\boldsymbol{i} + \frac{\partial w}{\partial y}\boldsymbol{j} + \frac{\partial w}{\partial z}\boldsymbol{k},
\end{aligned} \tag{1.9}$$

where the partials are evaluated at P.

Example 1.1.2 For an example of a surface, consider the surface $(2,\phi,z)$ in cylindrical coordinates. It is the infinite cylinder of radius 2 about the $z$-axis (see figure above). A normal to this surface at the point P = (2,0,0) is the gradient $\nabla\rho$, a vector starting at P and pointing outward along the positive $x$-axis.

Compare the coefficients of ***i***, ***j***, and ***k*** in the definitions (1.6) and (1.9). For example, $\frac{\partial x}{\partial u}$ and $\frac{\partial u}{\partial x}$. Are they are inverses of one another? If so, when $\frac{\partial x}{\partial u} = 0$ then $\frac{\partial u}{\partial x} = \frac{1}{0}$ would not exist. If fact, they are *not* inverses of each other. When $\frac{\partial x}{\partial u} = 0$, it is due to $x$ not changing when $u$ does. But, then, $\frac{\partial u}{\partial x} = 0$ because $u$ would not change when $x$ does. This is because the parameters in the numerators are *coordinate functions* and the parameters in the denominators are *coordinates*, as we now explain.

Definition A **coordinate function** is a function of a point P = ($x$, $y$, $z$) that returns the value of a coordinate. In $\mathbb{R}^3$ there are three coordinate functions: $x(\mathrm{P}) = x$, $y(\mathrm{P}) = y$, and $z(\mathrm{P}) = z$. When we write $w = x(\mathrm{P}) = w(x,y,z)$, then

$$\frac{\partial w}{\partial x} = \lim_{\Delta x \to 0} \frac{w(x+\Delta x,y,z)-w(x,y,z)}{\Delta x} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)-x}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1,$$

and

$$\frac{\partial w}{\partial y} = \lim_{\Delta y \to 0} \frac{w(x,y+\Delta y,z)-w(x,y,z)}{\Delta y} = \lim_{\Delta y \to 0} \frac{x-x}{\Delta y} = \lim_{\Delta y \to 0} \frac{0}{\Delta y} = 0.$$

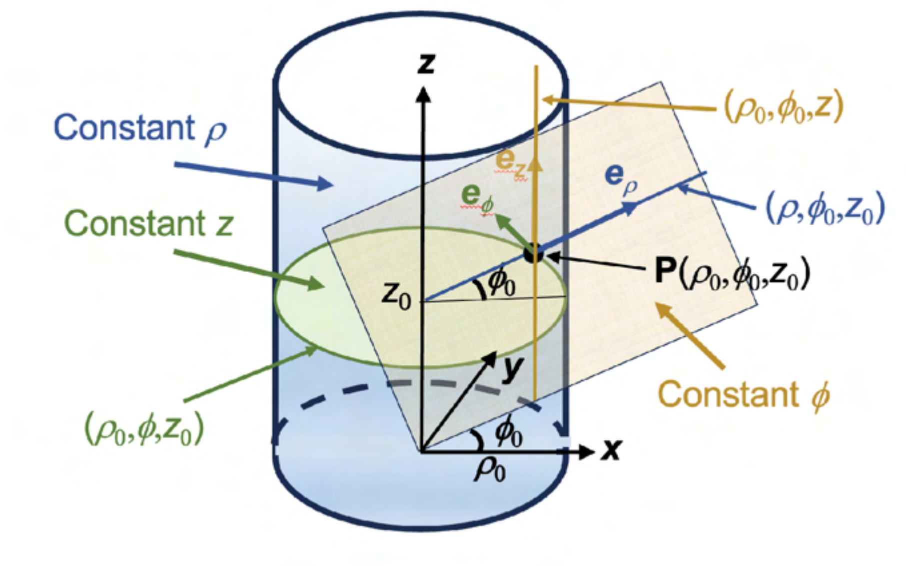
Observe that in the first expression, we are evaluating as the $x$-coordinate changes due to $\partial x$ in the denominator. Hence we write $w(x+\Delta x,y,z)$ in the numerator. In the second expression, we are evaluating as the $y$-coordinate changes due to $\partial y$ in the denominator. Hence we write $w(x,y+\Delta y,z)$ in the numerator.

Instead of $w$, we usually write the coordinate function using $x$ and $y$; that is, we write $x(\mathrm{P}) = x(x,y,z)$. Thus, with *partial* derivatives we have that $\frac{\partial x}{\partial x} = 1$ is short-hand for $\frac{\partial x(x,y,z)}{\partial x} = 1$, and $\frac{\partial x}{\partial y} = 0$ is short-hand for $\frac{\partial x(x,y,z)}{\partial y} = 0$, and we also have $\frac{\partial y}{\partial x} = 0$.

Just to be *very* clear, compare the derivative $\frac{dw}{dx}$ to the partial derivative $\frac{\partial w}{\partial x}$. The former expression applies when $w = w(x)$ is a function of just one variable. The latter applies when $w = w(x,y,z)$ is a function of several variables. Consider $y = x^2$, a function of just one variable. Then $x = \sqrt{y}$, and at the origin, $\left.\frac{dy}{dx}\right|_{(0,0)} = 2x\big|_{(0,0)} = 0$ but $\left.\frac{dx}{dy}\right|_{(0,0)} = \left.\frac{1}{2\sqrt{y}}\right|_{(0,0)}$ is undefined. With derivatives, $\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}$. We just saw that this relationship does not hold for partial derivatives.

We will see in the Corollary to Theorem 1.2.7 that the dual basis is an orthogonal basis iff the natural basis is an orthogonal basis; that the tangent vectors to the 1-parameter curves are mutually orthogonal iff the normal vectors to the 2-parameter surfaces are mutually orthogonal.

Exercise 1.1.3 Find the natural and dual bases for cylindrical coordinates.




The figure shows a cylinder, the surface of constant $\rho = \rho_0$, shaded in blue; the disc of constant $z = z_0$ in green; and the half-plane of constant $\phi = \phi_0$ in tan. The point P is at the intersection of the three surfaces. The boundary of the green disc is a circle having $\phi$ for its parameter, and, so, the natural basis vector $\mathbf{e}_\phi$ lies along a tangent to the circle at P. The blue line through P is parameterized by $\rho$, and, so $\mathbf{e}_\rho$ points outward from P as shown. It is on the blue line because the line is the tangent to itself. The vertical tan line through P is parameterized by $z$, and $\mathbf{e}_z$ points upward from P as shown, tangent to the line. The dual basis vectors (not shown) lie normal to the three surfaces and are parallel to their respective natural basis vectors.

The position vector is

$$\boldsymbol{r} = \boldsymbol{r}\,(\rho,\phi,z) \overset{(1.3)}{=} x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k} = \rho\, cos\phi\, \boldsymbol{i} + \rho\, sin\phi\, \boldsymbol{j} + z\,\boldsymbol{k},$$

where

$$x = \rho\, cos\phi \qquad y = \rho\, sin\phi \qquad z = z$$
$$\rho = \sqrt{x^2 + y^2} \qquad \phi = arctan\left[\tfrac{y}{x}\right] \qquad z = z\,.$$

The natural basis consists of the tangent vectors of $\boldsymbol{r}$ :

$$\mathbf{e}_\rho = \frac{\partial \boldsymbol{r}}{\partial \rho} = \frac{\partial x}{\partial \rho}\,\boldsymbol{i} + \frac{\partial y}{\partial \rho}\,\boldsymbol{j} + \frac{\partial z}{\partial \rho}\,\boldsymbol{k} = cos\phi\,\boldsymbol{i} + sin\phi\,\boldsymbol{j}$$
$$\mathbf{e}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = \frac{\partial x}{\partial \phi}\,\boldsymbol{i} + \frac{\partial y}{\partial \phi}\,\boldsymbol{j} + \frac{\partial z}{\partial \phi}\,\boldsymbol{k} = -\rho\, sin\phi\,\boldsymbol{i} + \rho\cos\phi\,\boldsymbol{j}$$
$$\mathbf{e}_z = \frac{\partial \boldsymbol{r}}{\partial z} = \frac{\partial x}{\partial z}\,\boldsymbol{i} + \frac{\partial y}{\partial z}\,\boldsymbol{j} + \frac{\partial z}{\partial z}\,\boldsymbol{k} = \boldsymbol{k}\,,$$

where $\rho \geq 0$ and $0 \leq \phi < 2\pi$. Note that points P on the $z$-axis are represented by multiple different coordinates P = $(\rho,0,z)$, but each coordinate uniquely defines the point. Cylindrical coordinates cover all points of Euclidean 3-space, $\mathbb{E}^3$.

To find the dual basis, first convert the partial derivatives from $x$, $y$, $z$ to $\rho$, $\phi$, and $z$:

$$\frac{\partial \rho}{\partial x} = \frac{x}{\sqrt{x^2+y^2}} = \frac{\rho\, cos\phi}{\sqrt{\rho^2}} = \cos\phi, \quad \frac{\partial \rho}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} = \frac{\rho\, sin\phi}{\sqrt{\rho^2}} = sin\phi, \quad \frac{\partial \rho}{\partial z} = 0$$
$$\frac{\partial \phi}{\partial x} = \frac{-y}{x^2+y^2} = \frac{-\rho\, sin\phi}{\rho^2} = \frac{-\sin\phi}{\rho}, \quad \frac{\partial \phi}{\partial y} = \frac{x}{x^2+y^2} = \frac{\rho\, cos\phi}{\rho^2} = \frac{cos\phi}{\rho}, \quad \frac{\partial \phi}{\partial z} = 0$$
$$\frac{\partial z}{\partial x} = 0, \qquad \frac{\partial z}{\partial y} = 0, \qquad \frac{\partial z}{\partial z} = 1$$

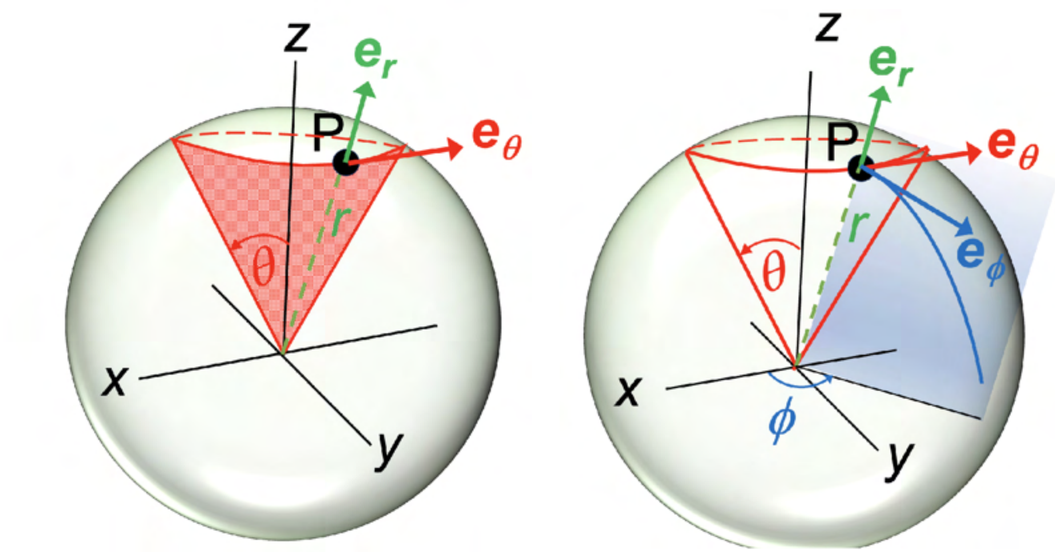
The dual basis is composed of the gradient vectors, defined in terms of the partials:

$$\mathbf{e}^\rho = \nabla\rho = \frac{\partial \rho}{\partial x}\,\boldsymbol{i} + \frac{\partial \rho}{\partial y}\,\boldsymbol{j} + \frac{\partial \rho}{\partial z}\,\boldsymbol{k} = cos\phi\,\boldsymbol{i} + sin\phi\,\boldsymbol{j}$$
$$\mathbf{e}^\phi = \nabla\phi = \frac{\partial \phi}{\partial x}\,\boldsymbol{i} + \frac{\partial \phi}{\partial y}\,\boldsymbol{j} + \frac{\partial \phi}{\partial z}\,\boldsymbol{k} = -\frac{sin\phi}{\rho}\,\boldsymbol{i} + \frac{\cos\phi}{\rho}\,\boldsymbol{j}$$
$$\mathbf{e}^z = \nabla z = \frac{\partial z}{\partial x}\,\boldsymbol{i} + \frac{\partial z}{\partial y}\,\boldsymbol{j} + \frac{\partial z}{\partial z}\,\boldsymbol{k} = \boldsymbol{k}$$

The dual basis is also defined at all points in $\mathbb{E}^3$.

Observe further that the corresponding natural and dual basis vectors have the same direction, only differing by up to a scalar factor. Further, observe that the basis vectors are orthogonal. For example, $\boldsymbol{e}_\rho \bullet \boldsymbol{e}_\phi = 0$. ∎

Example 1.1.4 Find the natural and dual bases for spherical coordinates




The RH figure shows three surfaces that intersect at a point P. The LH figure shows two of those surfaces, a green sphere of radius *r* about the origin and a red cone around the *z*-axis at angle $\theta$. The intersection is the circle passing through P. The basis vector $\boldsymbol{e}_\theta$ at point P is shown tangent to the circle. The basis vector $\boldsymbol{e}_r$ lies on the line that passes through the origin and P (because the tangent to a line is the line itself). The RH figure adds the half-plane generated by angle $\phi$, and the blue arc is where it intersects the sphere. $\boldsymbol{e}_\phi$ is shown tangent to that arc at P.

The position vector is

$$\boldsymbol{r} = \boldsymbol{r}(r,\theta,\phi) \overset{(1.3)}{=} x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k} = r\sin\theta\cos\phi\,\boldsymbol{i} + r\sin\theta\sin\phi\,\boldsymbol{j} + r\cos\theta\,\boldsymbol{k}$$

and

$$x = r\sin\theta\cos\phi \qquad y = r\sin\theta\sin\phi \qquad z = r\cos\theta$$

$$r = \sqrt{x^2+y^2+z^2} \qquad \theta = \arccos\frac{z}{r} \qquad \phi = \arctan\frac{y}{x}\,.$$

The natural basis is:

$$\boldsymbol{e}_1 = \boldsymbol{e}_r = \frac{\partial \boldsymbol{r}}{\partial r} = \frac{\partial x}{\partial r}\boldsymbol{i} + \frac{\partial y}{\partial r}\boldsymbol{j} + \frac{\partial z}{\partial r}\boldsymbol{k} = \sin\theta\cos\phi\,\boldsymbol{i} + \sin\theta\sin\phi\,\boldsymbol{j} + \cos\theta\,\boldsymbol{k}$$

$$\boldsymbol{e}_2 = \boldsymbol{e}_\theta = \frac{\partial \boldsymbol{r}}{\partial \theta} = \frac{\partial x}{\partial \theta}\boldsymbol{i} + \frac{\partial y}{\partial \theta}\boldsymbol{j} + \frac{\partial z}{\partial \theta}\boldsymbol{k} = r\cos\theta\cos\phi\,\boldsymbol{i} + r\cos\theta\sin\phi\,\boldsymbol{j} - r\sin\theta\,\boldsymbol{k}$$

$$\boldsymbol{e}_3 = \boldsymbol{e}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = \frac{\partial x}{\partial \phi}\boldsymbol{i} + \frac{\partial y}{\partial \phi}\boldsymbol{j} + \frac{\partial z}{\partial \phi}\boldsymbol{k} = -r\sin\theta\sin\phi\,\boldsymbol{i} + r\sin\theta\cos\phi\,\boldsymbol{j},$$

where $r \geq 0$ and $0 \leq \theta \leq \pi$ . As with cylindrical coordinates, spherical coordinates are defined at every point of $\mathbb{E}^3$ though points on the $z$-axis are each represented by the infinite number of coordinates $(0,\phi$

The dual basis is:

$$\boldsymbol{e}^1 = \boldsymbol{e}^r = \nabla \boldsymbol{r} = \frac{\partial r}{\partial x}\,\boldsymbol{i} + \frac{\partial r}{\partial y}\,\boldsymbol{j} + \frac{\partial r}{\partial z}\,\boldsymbol{k}$$

$$= \frac{x}{\sqrt{x^2+y^2+z^2}}\,\boldsymbol{i} + \frac{y}{\sqrt{x^2+y^2+z^2}}\,\boldsymbol{j} + \frac{z}{\sqrt{x^2+y^2+z^2}}\,\boldsymbol{k}$$

$$= \sin\theta\cos\phi\,\boldsymbol{i} + \sin\theta\sin\phi\,\boldsymbol{j} + \cos\theta\,\boldsymbol{k}$$

$$\boldsymbol{e}^2 = \boldsymbol{e}^\theta = \nabla\theta = \frac{\partial \theta}{\partial x}\,\boldsymbol{i} + \frac{\partial \theta}{\partial y}\,\boldsymbol{j} + \frac{\partial \theta}{\partial z}\,\boldsymbol{k}$$

$$= \frac{x\,z}{(x^2+y^2+z^2)\,\sqrt{x^2+y^2}}\,\boldsymbol{i} + \frac{y\,z}{(x^2+y^2+z^2)\,\sqrt{x^2+y^2}}\,\boldsymbol{j} - \frac{\sqrt{x^2+y^2}}{\sqrt{x^2+y^2+z^2}}\,\boldsymbol{k}$$

$$= \frac{\cos\theta\,\cos\phi}{r}\,\boldsymbol{i} + \frac{\cos\theta\,\sin\phi}{r}\,\boldsymbol{j} - \frac{\sin\theta}{r}\,\boldsymbol{k}$$

$$\boldsymbol{e}^3 = \boldsymbol{e}^\phi = \nabla\phi = \frac{\partial \phi}{\partial x}\,\boldsymbol{i} + \frac{\partial \phi}{\partial y}\,\boldsymbol{j} + \frac{\partial \phi}{\partial z}\,\boldsymbol{k} = -\frac{y}{x^2+y^2}\,\boldsymbol{i} + \frac{x}{x^2+y^2}\,\boldsymbol{j}$$

$$= -\frac{\sin\phi}{r\sin\theta}\,\boldsymbol{i} + \frac{\cos\phi}{r\sin\theta}\,\boldsymbol{j}\,. \qquad \blacksquare$$

## 1.2 Index Notation

Indices can be either subscripts or superscripts. This book also uses the term "suffix" to mean "index". $x^i$ denotes $x$, $y$, or $z$:

$$x^1 = x,\ \ x^2 = y,\ \text{ and }\ x^3 = z.$$

Similarly, for an alternate coordinate system $(u,v,w)$, $u^i$ denotes $u$, $v$, and $w$:

$$u^1 = u,\ \ u^2 = v,\ \text{ and }\ u^3 = w.$$

Coordinates, like $(x^i)$ and $(u^i)$, are always designated by superscripts, never by subscripts. Bases and vectors (like $\boldsymbol{\lambda}$, see below), on the other hand, come in two varieties, contravariant and covariant, and are expressed appropriately in terms of both subscripts and superscripts. The natural and dual basis vectors are denoted, respectively, as

$$\{\mathbf{e}_i\} = \{\mathbf{e}_u,\mathbf{e}_v,\mathbf{e}_w\} \qquad \text{and} \qquad \{\mathbf{e}^i\} = \{\mathbf{e}^u,\mathbf{e}^v,\mathbf{e}^w\}.$$

Using indices $i$ = 1, 2, 3, we can rewrite (1.6) and (1.9) as

$$\mathbf{e}_i \equiv \frac{\partial \boldsymbol{r}}{\partial u^i} = \frac{\partial x}{\partial u^i}\boldsymbol{i} + \frac{\partial y}{\partial u^i}\boldsymbol{j} + \frac{\partial z}{\partial u^i}\boldsymbol{k} \tag{1.2-1}$$

$$\mathbf{e}^i \equiv \nabla u^i = \frac{\partial u^i}{\partial x}\boldsymbol{i} + \frac{\partial u^i}{\partial y}\boldsymbol{j} + \frac{\partial u^i}{\partial z}\boldsymbol{k} \quad . \tag{1.2-2}$$

Convention  For the remainder of this book, indices $i$, $j$, $k$, ... will range over 1, 2, 3.

A vector $\boldsymbol{\lambda}$ can be expressed in terms of either contravariant or covariant basis vectors:

$$\boldsymbol{\lambda} = \lambda^u\mathbf{e}_u + \lambda^v\mathbf{e}_v + \lambda^w\mathbf{e}_w \quad \text{and} \quad \boldsymbol{\lambda} = \lambda_u\mathbf{e}^u + \lambda_v\mathbf{e}^v + \lambda_w\mathbf{e}^w\ .$$

We also express vectors in terms of their components:

$$\boldsymbol{\lambda} = (\lambda^u, \lambda^v, \lambda^w) \quad \text{and} \quad \boldsymbol{\lambda} = (\lambda_u, \lambda_v, \lambda_w).$$

Using loose terminology, we will often refer to $\boldsymbol{\lambda}$ as $\boldsymbol{\lambda} = (\lambda^i)$ or even $\boldsymbol{\lambda} = \lambda^i$, where

$$\lambda^1 = \lambda^u,\quad \lambda^2 = \lambda^v,\quad \text{and} \quad \lambda^3 = \lambda^w.$$

Einstein summation convention  When upper and lower index *variables* match,  it means to sum over those variables. For example,

$$\lambda^i\,\mathbf{e}_i \equiv \lambda^1\mathbf{e}_1 + \lambda^2\mathbf{e}_2 + \lambda^3\mathbf{e}_3\ .$$

To be used properly, an index variable may be used at most twice in any term, and then it must occur once as a subscript and once as a superscript. Using Einstein summation convention, we make the following definitions.

Definition $\lambda^i$ are called **contravariant** components and $\lambda_i$ are called **covariant** components. $\boldsymbol{\lambda} = \lambda^i \mathbf{e}_i$ is a **contravariant vector** and $\boldsymbol{\lambda} = \lambda_i \mathbf{e}^i$ is a **covariant vector**. $\mathbf{e}_i$ is a **natural basis vector** and $\mathbf{e}^i$ is a **dual basis vector**.

We observe in passing that $\mathbf{e}_i$ is a contravariant vector and $\mathbf{e}^i$ is covariant.

Given a vector $\boldsymbol{\lambda}$, it can be also be expressed in terms of the Cartesian basis:

$$\boldsymbol{\lambda} = \lambda^i \mathbf{e}_i \quad \text{or} \quad \boldsymbol{\lambda} = \lambda_i \mathbf{e}^i \quad \text{or} \quad \boldsymbol{\lambda} = \lambda(x)\, \boldsymbol{i} + \lambda(y)\, \boldsymbol{j} + \lambda(z)\, \boldsymbol{k} \tag{1.16}$$

Definition. $\lambda^i$ are called **contravariant** components and $\lambda_i$ are called **covariant** components. $\boldsymbol{\lambda} = \lambda^i \mathbf{e}_i$ is a **contravariant vector** and $\boldsymbol{\lambda} = \lambda_i \mathbf{e}^i$ is a **covariant vector**. $\mathbf{e}_i$ is a **natural basis vector** and $\mathbf{e}^j$ is a **dual basis vector**.

The **Kronecker delta** is $\delta^i_j \equiv \delta^j_i \equiv \delta_{ij} \equiv \delta_{ji} \equiv \delta^{ij} \equiv \delta^{ji} \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ (1.18)

The **dot product of two Cartesian vectors** $\boldsymbol{v} = v_1\, \boldsymbol{i} + v_2\, \boldsymbol{j} + v_3\, \boldsymbol{k}$ and $\boldsymbol{w} = w_1\, \boldsymbol{i} + w_2\, \boldsymbol{j} + w_3\, \boldsymbol{k}$ is $\boldsymbol{v} \bullet \boldsymbol{w} \equiv v_1\, w_1 + v_2\, w_2 + v_3\, w_3$

Theorem 1.2.1 $\mathbf{e}^i \bullet \mathbf{e}_j = \delta^i_j$ and $\mathbf{e}_i \bullet \mathbf{e}^j = \delta^j_i$. (1.17)

Proof. $\mathbf{e}^i \bullet \mathbf{e}_j \overset{(1.2-1,\,1.2-2)}{=} \left( \dfrac{\partial x}{\partial u^i} \boldsymbol{i} + \dfrac{\partial y}{\partial u^i} \boldsymbol{j} + \dfrac{\partial z}{\partial u^i} \boldsymbol{k} \right) \bullet \left( \dfrac{\partial u^i}{\partial x} \boldsymbol{i} + \dfrac{\partial u^i}{\partial y} \boldsymbol{j} + \dfrac{\partial u^i}{\partial z} \boldsymbol{k} \right)$

$$= \frac{\partial u^i}{\partial x} \frac{\partial x}{\partial u^j} + \frac{\partial u^i}{\partial y} \frac{\partial y}{\partial u^j} + \frac{\partial u^i}{\partial z} \frac{\partial z}{\partial u^j} \overset{\text{Chain Rule}}{=} \frac{\partial u^i}{\partial u^j} = \delta^i_j .$$

For example, $\dfrac{\partial u^2}{\partial u^2} = \dfrac{\partial v}{\partial v} = 1$ and $\dfrac{\partial u^2}{\partial u^3} = \dfrac{\partial v}{\partial w} = 0$ ■

The book does not mention this, but a corollary of Theorem 1.2.1 is that the natural and dual bases are in fact bases, which are sets of three linearly independent vectors.

Corollary $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and $\{\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3\}$ are linearly independent sets of vectors.

Proof. Suppose $a^i$ exist such that $a^i \mathbf{e}_i = 0$. Then each $a^i = 0$:

$$0 = 0 \bullet \mathbf{e}^j = (a^i \mathbf{e}_i) \bullet \mathbf{e}^j = a^i\, \delta^j_i = a^j \quad ■$$

The next theorem shows that the covariant basis picks out the contravariant components and the contravariant basis picks out the covariant components.

Theorem 1.2.2

$$\lambda^j = \boldsymbol{\lambda} \cdot \mathbf{e}^j \tag{1.19}$$

$$\lambda_j = \boldsymbol{\lambda} \cdot \mathbf{e}_j. \tag{1.20}$$

Proof. $\boldsymbol{\lambda} \cdot \mathbf{e}^j = \lambda^i\, \mathbf{e}_i \cdot \mathbf{e}^j \overset{1.17}{=} \lambda^i \;\; \delta_i^j = \lambda^j$ ■

If $\mathbf{e}^i \cdot \mathbf{e}_j \overset{(1.17)}{=} \delta_j^i$ and $\mathbf{e}_i \cdot \mathbf{e}^j \overset{(1.17)}{=} \delta_i^j$, then what about $\mathbf{e}_i \cdot \mathbf{e}_j$ and $\mathbf{e}^i \cdot \mathbf{e}^j$? For example,

$$\mathbf{e}_i \cdot \mathbf{e}_j = \left(\frac{\partial x}{\partial u^i}\boldsymbol{i} + \frac{\partial y}{\partial u^i}\boldsymbol{j} + \frac{\partial z}{\partial u^i}\boldsymbol{k}\right) \cdot \left(\frac{\partial x}{\partial u^j}\boldsymbol{i} + \frac{\partial y}{\partial u^j}\boldsymbol{j} + \frac{\partial z}{\partial u^j}\boldsymbol{k}\right) = \frac{(\partial x)^2 + (\partial y)^2 + (\partial z)^2}{\partial u^i\, \partial u^j}.$$

This does not further simplify (and even the above simplification assumes continuous 2nd partials in order for $\frac{\partial x}{\partial u^i}\frac{\partial y}{\partial u^j}$ to equal $\frac{\partial y}{\partial u^i}\frac{\partial x}{\partial u^j}$), so we assign symbols and names to these quantities and explore their properties.

Definition $\mathbf{e}_i \cdot \mathbf{e}_j$ and $\mathbf{e}^i \cdot \mathbf{e}^j$ are scalars and are denoted by

$$g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j \quad \text{and} \quad g^{ij} \equiv \mathbf{e}^i \cdot \mathbf{e}^j. \tag{1.22, 1.24}$$

$g_{ij}$ and $g^{ij}$ are called **metric tensors**. In Section 1.8, we will see that $g_{ij}$ and $g^{ij}$ satisfy the tensor definition (1.73). For now, “metric tensor” is just a label.

If $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are vectors, there are four ways to express their dot product, corresponding to the four combinations of subscripts and superscripts of $\mathbf{e}_{\square}^{\square} \cdot \mathbf{e}_{\square}^{\square}$. Two of them are Kronecker deltas and two are metric tensors.

Theorem 1.2.3 $\boldsymbol{\lambda} \cdot \boldsymbol{\mu} = \lambda_i\, \mu^i = \lambda^i\, \mu_i = g_{ij}\, \lambda^i\, \mu^j = g^{ij}\, \lambda_i\, \mu_j.$ (1.26)

Proof. $\boldsymbol{\lambda} \cdot \boldsymbol{\mu} = \lambda_i\, \mathbf{e}^i \cdot \mu^j\, \mathbf{e}_j = \lambda_i\, \mu^j\, \delta^i{}_j = \lambda_i\, \mu^i$ (up-down indices, Kronecker delta)

$\boldsymbol{\lambda} \cdot \boldsymbol{\mu} = \lambda^i\, \mathbf{e}_i \cdot \mu^j\, \mathbf{e}_j = g_{ij}\, \lambda^i\, \mu^j$ (down-down indices, metric tensor) ■

The next theorem and its corollary show that the metric tensors raise and lower indices.

Theorem 1.2.4 $g^{ij}\,\lambda_j = \lambda^i$ and $g_{ij}\,\lambda^j = \lambda_i$. (1.27 - 1.28)

Proof. $\forall\,\lambda_i,\ \lambda_i\,[g^{ij}\,\lambda_j] = g^{ij}\,\lambda_i\,\lambda_j \overset{1.26}{=} \lambda_i\,[\lambda^i] \Rightarrow g^{ij}\,\lambda_j = \lambda^i$ ■

Corollary (Exercise 1.2.2) $g^{ij}\,\mathbf{e}_j = \mathbf{e}^i$ and $g_{ij}\,\mathbf{e}^j = \mathbf{e}_i$. (1.2-3, 1.2-4)

Proof. $\forall\,\boldsymbol{\lambda},\ \boldsymbol{\lambda}\bullet(g^{ij}\,\mathbf{e}_j) = g^{ij}\,\boldsymbol{\lambda}\bullet\mathbf{e}_j \overset{1.19}{=} g^{ij}\lambda_j \overset{1.27}{=} \lambda^i \overset{1.19}{=} \boldsymbol{\lambda}\bullet\mathbf{e}^i \Rightarrow \mathbf{e}^i = g^{ij}\,\mathbf{e}_j$ ■

The following theorem shows that the metric tensors are inverses of each other. (See 1.33, below, for an additional display of this fact.)

Theorem 1.2.5 $g^{ij}\,g_{jk} = \delta^i_k$ and $g_{ij}\,g^{jk} = \delta^k_i$ . (1.29, 1.31)

Proof. $\forall\,\lambda^k,\ [g^{ij}\,g_{jk}]\,\lambda^k = g^{ij}\left[g_{jk}\,\lambda^k\right] \overset{(1.28)}{=} g^{ij}\,\lambda_j \overset{(1.27)}{=} \lambda^i \Rightarrow g^{ij}\,g_{jk} = \delta^i_k$ ■

Notation Vectors and matrices with contravariant and covariant indices:

**Vectors:** $L = \begin{pmatrix}\lambda^1\\ \lambda^2\\ \lambda^3\end{pmatrix}$, $\underset{\wedge}{L} = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix}$ **Matrices**: $\hat{M} = \left(m^{ij}\right)$, $M = (m_{ij})$

Definition The **metric tensors** have matrix representations $G = (g_{ij})$ and $\hat{G} = \left(g^{ij}\right)$. Index $i$ represents row $i$ and index $j$ represents column $j$ .

Theorem 1.2.6 (Exercise 1.2.4)

The natural basis is orthogonal iff $\exists$ scalars $k_{ij} \ni G = \begin{pmatrix}k_{11} & 0 & 0\\ 0 & k_{22} & 0\\ 0 & 0 & k_{33}\end{pmatrix}$.

The natural basis is orthonormal iff $G = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$.

The dual basis is orthogonal iff $\exists$ scalars $k^{ij} \ni \hat{G} = \begin{pmatrix}k^{11} & 0 & 0\\ 0 & k^{22} & 0\\ 0 & 0 & k^{33}\end{pmatrix}$.

The dual basis is orthonormal iff $\hat{G} = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$.

Proof. $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are mutually orthogonal iff $\exists$ scalars $k_{ij} \ni g_{ij} \overset{1.22}{\equiv} \mathbf{e}_i \cdot \mathbf{e}_j = k_{ij}\,\delta_{ij}$ . They are orthonormal iff $g_{ij} \overset{1.22}{\equiv} \mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ . The proof is similar for the dual basis. ■

Corollary The natural and dual bases are orthogonal iff they point in the same directions: $\mathbf{e}_1 = k_{11}\,\mathbf{e}^1$, $\mathbf{e}_2 = k_{22}\,\mathbf{e}^2$, and $\mathbf{e}_3 = k_{33}\,\mathbf{e}^3$. The natural and dual bases are orthonormal iff they coincide: $\mathbf{e}_1 = \mathbf{e}^1$, $\mathbf{e}_2 = \mathbf{e}^2$, and $\mathbf{e}_3 = \mathbf{e}^3$.

Proof: Natural basis is orthogonal

$$\Leftrightarrow \quad \exists\, k_{ij} \text{ such that } (g_{ij}) = G = \begin{pmatrix} k_{11} & 0 & 0 \\ 0 & k_{22} & 0 \\ 0 & 0 & k_{33} \end{pmatrix} \Leftrightarrow \mathbf{e}_i \overset{(1.2-4)}{=} g_{ij}\,\mathbf{e}^j = k_{ii}\,\mathbf{e}^i ,$$

and similarly for dual bases ■

Example Show that spherical $(r,\theta,\phi)$ and cylindrical $(\rho,\phi,z)$ coordinate systems have orthogonal bases by showing that their respective metric tensor matrices $G$ are diagonal. Show that $G$ and $\hat{G}$ are inverse matrices.

Spherical

Natural Basis (Computed in Example 1.1.4, earlier) For $r > 0$ and $\theta > 0$,

$$\begin{aligned} \mathbf{e}_1 &= \mathbf{e}_r = \sin\theta\cos\phi\,\boldsymbol{i} + \sin\theta\sin\phi\,\boldsymbol{j} + \cos\theta\,\boldsymbol{k} \\ \mathbf{e}_2 &= \mathbf{e}_\theta = r\cos\theta\cos\phi\,\boldsymbol{i} + r\cos\theta\sin\phi\,\boldsymbol{j} - r\sin\theta\,\boldsymbol{k} \\ \mathbf{e}_3 &= \mathbf{e}_\phi = -r\sin\theta\sin\phi\,\boldsymbol{i} + r\sin\theta\cos\phi\,\boldsymbol{j} \end{aligned}$$

$$\therefore \quad G = (g_{ij}) = (\mathbf{e}_i \cdot \mathbf{e}_j) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix}, \text{ a diagonal matrix.} \quad ✓$$

Dual Basis (Computed in Example 1.1.4, earlier) For $r > 0$ and $\theta > 0$,

$$\begin{aligned} \mathbf{e}^1 &= \mathbf{e}^r = \sin\theta\cos\phi\,\boldsymbol{i} + \sin\theta\sin\phi\,\boldsymbol{j} + \cos\theta\,\boldsymbol{k} \\ \mathbf{e}^2 &= \mathbf{e}^\theta = \frac{\cos\theta\cos\phi}{r}\,\boldsymbol{i} + \frac{\cos\theta\sin\phi}{r}\,\boldsymbol{j} - \frac{\sin\theta}{r}\,\boldsymbol{k} \\ \mathbf{e}^3 &= \mathbf{e}^\phi = -\frac{\sin\phi}{r\sin\theta}\,\boldsymbol{i} + \frac{\cos\phi}{r\sin\theta}\,\boldsymbol{j} \end{aligned}$$

$$\hat{G} = \left(g^{ij}\right) = (\mathbf{e}^i \bullet \mathbf{e}^j) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2\theta} \end{pmatrix} \text{, a diagonal matrix.} \quad ✓$$

$G\,\hat{G} = I$ ✓ (See, also, 1.33, below.)

Note: The metric components $g_{ij}$ are all finite, but the components $g^{22}$ and $g^{33}$ blow up at points in which $r = 0$.

Cylindrical

<u>Natural Basis</u> (Computed in Exercise 1.1.3, earlier) For $\rho > 0$,

$\mathbf{e}_1 = \mathbf{e}_\rho = \cos\phi\ \boldsymbol{i} + \sin\phi\ \boldsymbol{j}$ $\mathbf{e}_2 = \mathbf{e}_\phi = -\rho \sin\phi\ \boldsymbol{i} + \rho\cos\phi\ \boldsymbol{j}$ $\mathbf{e}_3 = \mathbf{e}_z = \boldsymbol{k}$

$$\therefore\ G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{, a diagonal matrix} \quad ✓$$

<u>Dual Basis</u> (Computed in Exercise 1.1.3, earlier) For $\rho > 0$,

$$\mathbf{e}^1 = \mathbf{e}^r = \sin\theta\cos\phi\ \boldsymbol{i} + \sin\theta\sin\phi\ \boldsymbol{j} + \cos\theta\ \boldsymbol{k}$$

$$\mathbf{e}^2 = \mathbf{e}^\theta = \frac{\cos\theta\cos\phi}{r}\ \boldsymbol{i} + \frac{\cos\theta\sin\phi}{r}\ \boldsymbol{j} - \frac{\sin\theta}{r}\ \boldsymbol{k}$$

$$\mathbf{e}^3 = \mathbf{e}^\phi = -\frac{\sin\phi}{r\sin\theta}\ \boldsymbol{i} + \frac{\cos\phi}{r\sin\theta}\ \boldsymbol{j}$$

$$\hat{G} = \left(g^{ij}\right) = (\mathbf{e}^i \bullet \mathbf{e}^j) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{, a diagonal matrix} \quad ✓$$

$G\,\hat{G} = I$ ✓ ■

Notation Recall that $L = (\lambda^i)$ and $\underset{\wedge}{L} = (\lambda_i)$ denote column vectors. **Row vectors** are denoted with a transpose symbol: $L^{\mathrm{T}} = (\lambda^1, \lambda^2, \lambda^3)$ and $\underset{\wedge}{L}^{\mathrm{T}} = (\lambda_1, \lambda_2, \lambda_3)$. The **transpose of a matrix** is denoted $G^{\mathrm{T}}$. The **identity matrix** is often loosely denoted by a typical element, $\delta^i_j = I$. Another common practice that saves space when working only with column vectors is to loosely write $\lambda^i = \begin{pmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{pmatrix}$ and $\lambda_i = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$.

**Theorem 1.2.7** For any natural basis and corresponding dual basis,

$$\hat{G}\, G = I \;\text{ and }\; G\, \hat{G} = I \;\Leftrightarrow\; \boxed{\hat{G} = G^{-1}} \tag{1.33}$$

$$G\, L = \underset{\wedge}{L} \;\text{ and }\; G^{-1} \underset{\wedge}{L} = L \quad \text{(i.e., G and } G^{-1} \text{ raise and lower indices)} \tag{1.34}$$

$$g_{ij} = g_{ji} \;\text{ and }\; g^{ij} = g^{ji} \tag{1.30}$$

$$G = G^{\mathrm{T}} \;\text{ and }\; G^{-1} = \left(G^{-1}\right)^{\mathrm{T}} \tag{1.32}$$

Proof. $\hat{G}\, G \overset{(1.29)}{=} I,\;\; G\, \hat{G} \overset{(1.31)}{=} I,\;\; G\, L \overset{(1.28)}{=} \underset{\wedge}{L},\;\; G^{-1} \underset{\wedge}{L} \overset{(1.27)}{=} L,$

$$g_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j = \mathbf{e}_j \bullet \mathbf{e}_i = g_{ji}\,,\;\; G \overset{(1.30)}{=} G^{\mathbf{T}} \;\text{ and }\; G^{-1} \overset{(1.30)}{=} \left(G^{-1}\right)^{\mathrm{T}} \qquad \blacksquare$$

**Corollary** The natural basis is orthogonal iff the dual basis is orthogonal.

Proof. If $G = \begin{pmatrix} k_{11} & 0 & 0 \\ 0 & k_{22} & 0 \\ 0 & 0 & k_{33} \end{pmatrix}$ then $G^{-1} = \begin{pmatrix} 1/k_{11} & 0 & 0 \\ 0 & 1/k_{22} & 0 \\ 0 & 0 & 1/k_{33} \end{pmatrix}$.

So, the natural basis is orthogonal

$$\overset{\text{(Th 1.26)}}{\Longleftrightarrow} \quad G = \begin{pmatrix} k_{11} & 0 & 0 \\ 0 & k_{22} & 0 \\ 0 & 0 & k_{33} \end{pmatrix} \Longleftrightarrow \hat{G} = G^{-1} = \begin{pmatrix} 1/k_{11} & 0 & 0 \\ 0 & 1/k_{22} & 0 \\ 0 & 0 & 1/k_{33} \end{pmatrix}$$

$\overset{\text{(Th 1.26)}}{\Longleftrightarrow}$ The dual basis is orthogonal. $\blacksquare$

**Theorem 1.2.8** Letting $M = (\mu^i)$, the four ways of expressing $\boldsymbol{\lambda} \bullet \boldsymbol{\mu}$ in (1.26) can now be re-expressed in matrix form:

$$\boldsymbol{\lambda} \bullet \boldsymbol{\mu} = \underset{\wedge}{L}^{\mathrm{T}} M = L^{\mathrm{T}} \underset{\wedge}{M} = L^{\mathrm{T}} G\, M = \underset{\wedge}{L}^{\mathrm{T}}\, G^{-1}\, \underset{\wedge}{M}. \tag{1.35}$$

Since both spherical and cylindrical coordinates have orthogonal bases, by the corollary (above) to Exercise 1.2.4 we know that $\mathbf{e}_i$ and $\mathbf{e}^i$ differ only in magnitude. The next exercise provides a more interesting example where the bases are not orthogonal.

**Examples 1.1.3 and 1.2.1** Define a coordinate system $(u,v,w)$ by

$$x = u + v,\;\; y = u - v,\;\; z = 2uv + w \qquad \text{where } -\infty < u,\, v,\, w < \infty. \tag{1.10}$$

**Ex 1.1.3)** Find the natural and dual bases.

**Ex 1.2.1a)** Find the metric tensor matrices $G$ and $G^{-1}$ and verify that

$$G\, G^{-1} = G^{-1} G = I.$$

Ex 1.2.1b) Let the vector $\boldsymbol{\lambda} = \boldsymbol{i}$ have column vectors $L = (\lambda^k)$ and $\underset{\wedge}{L} = (\lambda_k)$.
Confirm equations (1.34) & (1.35): $\underset{\wedge}{L} = G\,L$ and $\boldsymbol{\lambda} \bullet \boldsymbol{\lambda} = L^T G\,L = L^T \underset{\wedge}{L}$.

<u>Ex 1.1.3</u>:

Inverting these equations yields

$$u = \tfrac{1}{2}(x+y),\ \ v = \tfrac{1}{2}(x-y),\ \ w = z - \tfrac{1}{2}\left(x^2-y^2\right) \qquad (1.11)$$

Observe that the surface $u = u_0$ is a plane since it is a linear function of $v$ and $w$:

$$u = u_0 \Rightarrow x = u_0 + v,\quad y = u_0 - v,\quad z = 2u_0 v + w.$$

Similarly the surface $v = v_0$ is a plane. However, the surface $w = w_0$ is the hyperbolic paraboloid $z = 2uv + w_0$.

The position vector is $\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k} = (u+v)\,\boldsymbol{i} + (u-v)\,\boldsymbol{j} + (2uv+w)\,\boldsymbol{k}$.

The natural and dual bases are

$$\begin{aligned}
\boldsymbol{e}_1 &= \frac{\partial \boldsymbol{r}}{\partial u} = \boldsymbol{i} + \boldsymbol{j} + 2\,v\,\boldsymbol{k} & \qquad \boldsymbol{e}^1 &= \nabla u = \tfrac{1}{2}\,\boldsymbol{i} + \tfrac{1}{2}\,\boldsymbol{j} \\
\boldsymbol{e}_2 &= \frac{\partial \boldsymbol{r}}{\partial v} = \boldsymbol{i} - \boldsymbol{j} + 2\,u\,\boldsymbol{k} \qquad \text{and} & \qquad \boldsymbol{e}^2 &= \nabla v = \tfrac{1}{2}\,\boldsymbol{i} - \tfrac{1}{2}\,\boldsymbol{j} \\
\boldsymbol{e}_3 &= \frac{\partial \boldsymbol{r}}{\partial w} = \boldsymbol{k} & \qquad \boldsymbol{e}^3 &= \nabla w = -x\,\boldsymbol{i} + y\,\boldsymbol{j} + \boldsymbol{k} \\
& & &= -(u+v)\,\boldsymbol{i} + (u-v)\,\boldsymbol{j} + \boldsymbol{k}
\end{aligned}$$

Note that the basis elements are neither unit vectors (except for $\boldsymbol{e}_3$) nor orthogonal. For example, $\mathbf{e}_1 \bullet \mathbf{e}_2 = 4\,uv$. Also, $\boldsymbol{e}_1$ and $\mathbf{e}^1$ do not have the same direction. However, $\boldsymbol{e}_i \bullet \boldsymbol{e}^j = \delta_i^j$, confirming that (1.17) holds.

<u>Ex 1.2.1a</u>**:**

$$G = (g_{ij}) = (\mathbf{e}_i \bullet \mathbf{e}_j) = \begin{pmatrix} 2\left(1+2\,v^2\right) & 4\,uv & 2\,v \\ 4\,uv & 2\left(1+2\,u^2\right) & 2\,u \\ 2\,v & 2\,u & 1 \end{pmatrix}$$

$$G^{-1} = \left(g^{ij}\right) = \left(\mathbf{e}^i \bullet \mathbf{e}^j\right) = \begin{pmatrix} \frac{1}{2} & 0 & -v \\ 0 & \frac{1}{2} & -u \\ -v & -u & 2\,u^2+2\,v^2+1 \end{pmatrix}$$

$$G\,G^{-1} = G^{-1}\,G = I. \quad ✔$$

Ex 1.2.1b:

$$L = \left(\lambda^k\right) \overset{(1.19)}{=} \left(\lambda \bullet \boldsymbol{e}^k\right) = \left(\boldsymbol{i} \bullet \boldsymbol{e}^k\right) = \begin{pmatrix} \boldsymbol{i} \bullet \boldsymbol{e}^1 \\ \boldsymbol{i} \bullet \boldsymbol{e}^2 \\ \boldsymbol{i} \bullet \boldsymbol{e}^3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -(u+v) \end{pmatrix},$$

$$\mathrm{L}^{\mathrm{T}} = \left(\tfrac{1}{2} \;\; \tfrac{1}{2} \;\; -u - v\right), \text{ and } \underset{\wedge}{L} = (\lambda_i) \overset{(1.20)}{=} (\boldsymbol{\lambda} \bullet \boldsymbol{e}_i) = (\boldsymbol{i} \bullet \boldsymbol{e}_i) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = G\,L\,. \quad ✔$$

Also,

$$\boldsymbol{\lambda} \bullet \boldsymbol{\lambda} = \boldsymbol{i} \bullet \boldsymbol{i} = 1, \quad L^{\mathrm{T}} \mathrm{G}\, L = L^{\mathrm{T}} \underset{\wedge}{L} = 1 \quad ✔ \quad ■$$

## 1.3 Tangents and Gradients

Relaxing the requirement for bases to be orthonormal has led to development of two basis sets, $\{\mathbf{e}_i\}$ and $\{\mathbf{e}^i\}$ for Euclidean 3-space. We show below that both bases are useful.

If $u$, $v$, and $w$ are differential functions of $t$ on some interval $\mathrm{I} = [a,b]$, then $x = x(\, u(t),v(t),w(t)\,) = x(t)$, $y = y(t)$, and $z = z(t)$ are also differential functions of $t$ on I. Thus, $\boldsymbol{r}(t) = x(t)\,\boldsymbol{i} + y(t)\,\boldsymbol{j} + z(t)\,\boldsymbol{k} = x(u,v,w)\,\boldsymbol{i} + y(u,v,w)\,\boldsymbol{j} + z(u,v,w)\,\boldsymbol{k}$ traces out a curve $\gamma$, and $\dfrac{d\mathbf{r}}{dt}$ is tangent to $\gamma$ at each point $(u, v, w)$.

$$\dot{\boldsymbol{r}}\,(t) = \frac{d\mathbf{r}}{dt} = \frac{\partial\mathbf{r}}{du}\frac{du}{dt} + \frac{\partial\mathbf{r}}{dv}\frac{dv}{dt} + \frac{\partial\mathbf{r}}{dw}\frac{dw}{dt}$$
$$\overset{(1.2-1)}{=} \dot{u}(t)\,\mathbf{e}_u + \dot{v}(t)\,\mathbf{e}_v + \dot{w}(t)\,\mathbf{e}_w$$

Setting $(u^i) = (u,v,w)$, then $u^i$ is a differential function of $t$, and we have

$$\boxed{\dot{\boldsymbol{r}}\,(t) = \dot{u}^i\,(t)\,\mathbf{e}_i} \qquad (1.37)$$

Equation (1.37) shows that to investigate tangents to a space curve $\gamma$, the natural basis is the appropriate choice. Equation (1.37) can also be used to generate arc length. First, observe that the length L of a vector $\boldsymbol{v} = a\,\boldsymbol{i} + b\,\boldsymbol{j} + c\,\boldsymbol{k}$ is

$$L^2 = a^2 + b^2 + c^2 = \boldsymbol{v} \bullet \boldsymbol{v}.$$

Definition The length d$s$ of an infinitesimal portion of a curve is called the **line element**, calculated as

$$\boxed{ds^2 = \dot{\boldsymbol{r}} \bullet \dot{\boldsymbol{r}}\,dt^2 \overset{(1.37)}{=} g_{ij}\,\dot{u}^i\,\dot{u}^j\,dt^2} \qquad (1.39)$$

because $g_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j$. Equation (1.39) generalizes the Cartesian arc length, $ds^2 = dx^2 + dy^2 + dz^2$. The length of the curve $\gamma$ is

$$\boxed{L = \int_\gamma ds = \int_a^b \sqrt{g_{ij}\,\dot{u}^i\,\dot{u}^j}\,dt}. \qquad (1.38)$$

Since $\dot{u}^i = \dfrac{du^i}{dt}$, then $du^i = \dot{u}^i dt$, and we can also express equation (1.39) as

$$\boxed{ds^2 = g_{ij}\,du^i\,du^j} \qquad (1.3\text{-}1)$$

We next show that the dual basis $\{\mathbf{e}^i\}$ is the appropriate choice for investigating gradients of surfaces. Suppose that $\varphi$ is a differentiable function of $(u,v,w)$. Then $\varphi$ is also a differentiable function of $(x,y,z)$. That is, we can write

$$\varphi(u,v,w) = \varphi(\, u(x,y,z),\, v(x,y,z),\, w(x,y,z)\,)\,.$$

The gradient of $\varphi$ is

$$\nabla\varphi = \frac{\partial\varphi}{\partial x}\boldsymbol{i} + \frac{\partial\varphi}{\partial y}\boldsymbol{j} + \frac{\partial\varphi}{\partial z}\boldsymbol{k}. \tag{1.3-2}$$

Since $\dfrac{\partial\varphi}{\partial x} = \dfrac{\partial\varphi}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial\varphi}{\partial v}\dfrac{\partial v}{\partial x} + \dfrac{\partial\varphi}{\partial w}\dfrac{\partial w}{\partial x}$ and similarly for $\dfrac{\partial\varphi}{\partial y}$ and $\dfrac{\partial\varphi}{\partial z}$,

$$\nabla\varphi = \frac{\partial\varphi}{\partial u}\left(\frac{\partial u}{\partial x}\boldsymbol{i} + \frac{\partial u}{\partial y}\boldsymbol{j} + \frac{\partial u}{\partial z}\boldsymbol{k}\right) + \frac{\partial\varphi}{\partial v}\left(\frac{\partial v}{\partial x}\boldsymbol{i} + \frac{\partial v}{\partial y}\boldsymbol{j} + \frac{\partial v}{\partial z}\boldsymbol{k}\right)$$
$$+ \frac{\partial\varphi}{\partial w}\left(\frac{\partial w}{\partial x}\boldsymbol{i} + \frac{\partial w}{\partial y}\boldsymbol{j} + \frac{\partial w}{\partial z}\boldsymbol{k}\right)$$
$$\overset{(1.3-2)}{=} \frac{\partial\varphi}{\partial u}\nabla u + \frac{\partial\varphi}{\partial v}\nabla v + \frac{\partial\varphi}{\partial w}\nabla w \overset{(1.9)}{=} \frac{\partial\varphi}{\partial u}\mathbf{e}^u + \frac{\partial\varphi}{\partial v}\mathbf{e}^v + \frac{\partial\varphi}{\partial w}\mathbf{e}^w$$

$$\boxed{\nabla\varphi = \frac{\partial\varphi}{\partial u^i}\,\mathbf{e}^i \equiv \partial_i\varphi\,\mathbf{e}^i \equiv \varphi_{,i}\,\mathbf{e}^i} \tag{1.41 - 1.42}$$

Equations (1.41) and (1.42) reflect the fact that the $u^i$ superscript in the denominator is regarded as a subscript for the purpose of Einstein summation.

## 1.4 Coordinate Transformations in Euclidean 3-Space

Notation  Points P = ($u$,$v$,$w$) are also expressed as P = ($u^i$). We only use superscripts for coordinates, never subscripts.

Suppose a curved surface in space has two overlapping Euclidean coordinate systems and we wish to express each system in terms of the other. Let $\{u^i\} = \{u,v,w\}$ be an unprimed coordinate system centered at a point P and $\{u^{i'}\} = \{u', v', w'\}$ a primed coordinate systems centered at a point Q (possibly equal to P). Note that the primes have been put on the indices, not the bases. Some books put the primes on the bases.

Notation

Primed vectors: $\overset{'}{L} = (\lambda^{i'})$ and $\underset{'}{L} = (\lambda_{i'})$

[Compare to $L = (\lambda^i)$ and $\underset{\wedge}{L} = (\lambda_i)$]

Primed matrices: $\overset{'}{U} = (U_j^{i'})$ and $\underset{'}{U} = (U_{i'}^{j})$

In this section we will obtain a transformation matrix $\overset{'}{U}$ directly from the representation of the unprimed natural bases $\mathbf{e}_i$ in terms of the primed natural bases $\mathbf{e}_{j'}$. Then we will obtain another transformation matrix, $\underset{'}{U}$, from the representation of the unprimed dual bases $\boldsymbol{e}^i$ in terms of the primed dual bases $\boldsymbol{e}^{j'}$, and we will show that it equals the inverse matrix of $\overset{'}{U}$.

We can express equation (1.3) in terms of the primed and unprimed systems:

$$\boldsymbol{r} = x(u^i)\,\boldsymbol{i} + y(u^i)\,\boldsymbol{j} + z(u^i)\,\boldsymbol{k} = x(u^{i'})\,\boldsymbol{i} + y(u^{i'})\,\boldsymbol{j} + z(u^{i'})\,\boldsymbol{k} \tag{1.4-1}$$

We can rewrite (1.2-1) as

$$\mathbf{e}_j = \frac{\partial \boldsymbol{r}}{\partial u^j} \tag{1.4-2}$$

$$\mathbf{e}_{i'} = \frac{\partial \boldsymbol{r}}{\partial u^{i'}} \tag{1.4-3}$$

$$\therefore \quad \mathbf{e}_j \overset{(1.4-2)}{=} \frac{\partial \boldsymbol{r}}{\partial u^j} \overset{\text{(Chain Rule)}}{=} \frac{\partial \boldsymbol{r}}{\partial u^{i'}} \frac{\partial u^{i'}}{\partial u^j} \overset{(1.4-3)}{=} \frac{\partial u^{i'}}{\partial u^j}\, \mathbf{e}_{i'}\,. \tag{1.4-4}$$

$\dfrac{\partial u^{i'}}{\partial u^j}$ can be considered to be the $(i', j)$ element, $U_j^{i'}$, of a matrix $\acute{U}$. That is,

$$U_j^{i'} \equiv \frac{\partial u^{i'}}{\partial u^j} \quad \text{and} \quad \acute{U} \equiv (U_j^{i'}).$$

Therefore,

$$\mathbf{e}_j \overset{(1.4\text{–}4)}{=} U_j^{i'} \mathbf{e}_{i'} \;. \tag{1.44}$$

Because the primed and unprimed systems are on equal footing, we can also express this as

$$\mathbf{e}_{j'} \overset{(1.44)}{=} U_{j'}^{i} \mathbf{e}_i \;, \tag{1.4-5}$$

where

$$U_{j'}^{i} \equiv \frac{\partial u^{i}}{\partial u^{j'}} \quad \text{and} \quad \underset{'}{U} \equiv (U_{j'}^{i}).$$

Any vector, $\boldsymbol{\lambda}$, can be expressed in terms of basis vectors. So, we can write

$$\boldsymbol{\lambda} = \lambda^{i'} \mathbf{e}_{i'} \;. \tag{1.43}$$

Claim $\boxed{\lambda^{i'} = U_j^{i'} \lambda^j}$ : (1.45)

$$\lambda^{i'} \mathbf{e}_{i'} \overset{(1.43)}{=} \boldsymbol{\lambda} \overset{(1.16)}{=} \lambda^j \mathbf{e}_j \overset{(1.44)}{=} \lambda^j U_j^{i'} \mathbf{e}_{i'} \;\Rightarrow\; \text{for all } i',\; \lambda^{i'} = \lambda^j U_j^{i'} = U_j^{i'} \lambda^j \quad ✔$$

In matrix form this becomes $\quad \overset{'}{L} = \overset{'}{U} L \;.$ (1.4-6)

Observe that the primed basis vectors are on RHS of (1.44) but the primed components are on LHS of (1.45). We focus more on the coefficients of vectors, equation (1.45), than on the bases, equation (1.44). The coefficients represent a vector's magnitude while the bases represent its direction. However, we will show later that direction can be obtained from the metric tensor coefficients, eliminating the need for bases.

The dual basis transformation $\underset{\prime}{U}$ is developed similarly. From (1.2-2) we get

$$\mathbf{e}^j = \nabla u^j = \frac{\partial u^j}{\partial x}\boldsymbol{i} + \frac{\partial u^j}{\partial y}\boldsymbol{j} + \frac{\partial u^j}{\partial z}\boldsymbol{k} \tag{1.4-7}$$

$$\mathbf{e}^{i'} = \nabla u^{i'} = \frac{\partial u^{i'}}{\partial x}\boldsymbol{i} + \frac{\partial u^{i'}}{\partial y}\boldsymbol{j} + \frac{\partial u^{i'}}{\partial z}\boldsymbol{k} \tag{1.4-8}$$

$$\therefore \quad \mathbf{e}^j = \frac{\partial u^j}{\partial u^{i'}}\,\mathbf{e}^{i'}: \tag{1.4-9}$$

$$\mathbf{e}^j \overset{(1.4-7)}{=} \frac{\partial u^j}{\partial x}\boldsymbol{i} + \frac{\partial u^j}{\partial y}\boldsymbol{j} + \frac{\partial u^j}{\partial z}\boldsymbol{k}$$

$$\overset{\text{Chain Rule}}{=} \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial x}\boldsymbol{i} + \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial y}\boldsymbol{j} + \frac{\partial u^j}{\partial u^{i'}}\frac{\partial u^{i'}}{\partial z}\boldsymbol{k} \overset{(1.4-8)}{=} \frac{\partial u^j}{\partial u^{i'}}\,\mathbf{e}^{i'} \quad \checkmark$$

Using $U^j_{i'} \equiv \dfrac{\partial u^j}{\partial u^{i'}}$ (consistent with the above definition of $U^i_{j'}$) yields

$$\mathbf{e}^j \overset{(1.4-9)}{=} U^j_{i'}\,\mathbf{e}^{i'} \tag{1.46}$$

As in (1.16), we can express a vector, $\boldsymbol{\lambda}$, as a linear sum of primed basis vectors:

$$\boldsymbol{\lambda} = \lambda_{i'}\,\mathbf{e}^{i'} \tag{1.4-10}$$

Claim $\boxed{\lambda_{i'} = U^j_{i'}\,\lambda_j}$ : (1.47)

$$\lambda_{i'}\,\mathbf{e}^{i'} \overset{(1.4-2)}{=} \boldsymbol{\lambda} \overset{(1.16)}{=} \lambda_j\,\mathbf{e}^j \overset{(1.46)}{=} \lambda_j\,U^j_{i'}\,\mathbf{e}^{i'} \quad \Rightarrow \quad \text{for all } i',\ \lambda_{i'} = \lambda_j\,U^j_{i'} = U^j_{i'}\,\lambda_j \quad \checkmark$$

In matrix form this becomes

$$\underset{\prime}{L} = \underset{\prime}{U}\,\underset{\wedge}{L}\,. \tag{1.4-11}$$

Because the primed and unprimed coordinate systems are on equal footing, we can get the inverses by swapping indices:

$$\lambda^i \overset{(1.45)}{=} U^i_{j'}\,\lambda^{j'} \quad \text{and} \quad \lambda_i \overset{(1.47)}{=} U^{j'}_i\,\lambda_{j'} \tag{1.49}$$

or

$$L = \underset{\prime}{U}\,\overset{\prime}{L} \ \text{ and } \ \underset{\wedge}{L} = \overset{\prime}{U}\,\underset{\prime}{L}. \tag{1.4-12}$$

Solving the matrix equation (1.4-6), $\overset{\prime}{L} = \overset{\prime}{U}\,L$, for $L$ yields $L = \overset{\prime}{U}^{-1}\,\overset{\prime}{L}$. Substituting the expression for $L$ in equation (1.4-12) yields

$$\boxed{\underset{,}{U} = \overset{,}{U}^{-1}} \tag{1.4-13}$$

Consequently, we can greatly simplify our notation by setting

$$U = \overset{,}{U} = \left(\frac{\partial u^{i'}}{\partial u^j}\right) = (U^{i'}_j) \text{. Then } U^{-1} = \overset{,}{U}^{-1} \overset{(1.4-13)}{=} \underset{,}{U} = \left(\frac{\partial u^j}{\partial u^{i'}}\right) = (U^j_{i'}) \text{.}$$

So,

$$U^{-1}U = I = U\,U^{-1} \quad \text{or} \quad \boxed{U^k_{i'}\,U^{i'}_j = \delta^k_j = \delta^{k'}_{j'} = U^{k'}_i\,U^i_{j'}}\,. \tag{1.50}$$

Matrix equations (1.4-6) and (147b), above, can also now be simplified.

$$\overset{,}{L} = U\,L \quad \text{and, equivalently, } L = U^{-1}\overset{,}{L} \tag{1.4-14}$$

$$\underset{,}{L} = U^{-1}\underset{\wedge}{L} \quad \text{and, equivalently, } \underset{\wedge}{L} = U\,\underset{,}{L} \tag{1.4-15}$$

Definition The transformation matrices $\boxed{U = \left(\frac{\partial u^{i'}}{\partial u^j}\right)}$ and $\boxed{U^{-1} = \left(\frac{\partial u^j}{\partial u^{i'}}\right)}$ are known as **Jacobian matrices**. Their determinants, det $U$ and *det* $U^{-1}$, are called **Jacobians**. Since $U$ is invertible, det $U \neq 0 \neq$ *det* $U^{-1}$.

Example 1.4.1 Find the transformation matrices from spherical to cylindrical coordinates.

Let $(u^1,u^2,u^3) = (r,\theta,\phi)$ represent spherical coordinates as the unprimed system

Let $(u^{1'},u^{2'},u^{3'}) = (\rho,\phi,z)$ represent cylindrical coordinates as the primed system

$$\begin{cases} u^{1'} = \rho = r\sin\theta = u^1 \sin u^2 \\ u^{2'} = \phi = u^3 \\ u^{3'} = z = r\cos\theta = u^1\cos u^2 \end{cases}$$

$$\Rightarrow \quad U = \left(\frac{\partial u^{i'}}{\partial u^j}\right) = \begin{pmatrix} \frac{\partial u^{1'}}{\partial u^1} & \frac{\partial u^{1'}}{\partial u^2} & \frac{\partial u^{1'}}{\partial u^3} \\ \frac{\partial u^{2'}}{\partial u^1} & \frac{\partial u^{2'}}{\partial u^2} & \frac{\partial u^{2'}}{\partial u^3} \\ \frac{\partial u^{3'}}{\partial u^1} & \frac{\partial u^{3'}}{\partial u^2} & \frac{\partial u^{3'}}{\partial u^3} \end{pmatrix} = \begin{pmatrix} \sin u^2 & u^1\cos u^2 & 0 \\ 0 & 0 & 1 \\ \cos u^2 & -u^1\sin u^2 & 0 \end{pmatrix} = \begin{pmatrix} \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

$$\begin{cases} u^1 = r = \sqrt{\rho^2 + z^2} = \sqrt{\left(u^{1'}\right)^2 + \left(u^{3'}\right)^2} \\ u^2 = \theta = \arctan \dfrac{\rho}{z} = \arctan \dfrac{u^{1'}}{u^{3'}} \\ u^3 = \phi = u^{2'} \end{cases}$$

$$\Rightarrow \quad U^{-1} = \left(\frac{\partial u^i}{\partial u^{j'}}\right) = \begin{pmatrix} \frac{\rho}{\sqrt{\rho^2+z^2}} & 0 & \frac{z}{\sqrt{\rho^2+z^2}} \\ \frac{z}{\rho^2+z^2} & 0 & -\frac{\rho}{\rho^2+z^2} \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \sin\theta & 0 & \cos\theta \\ \frac{1}{r}\cos\theta & 0 & -\frac{1}{r}\sin\theta \\ 0 & 1 & 0 \end{pmatrix}$$

Check: $U\,U^{-1} = I$ ■

Let P = $(x,y,z) = (x^1, x^2, x^3) = (x^i)$. We sometimes treat P as a point and sometimes as a vector. But, is $x^i$ a vector in the sense of equation (1.45), $\lambda^{i'} = U_j^{i'}\,\lambda^j$ ? That is, does $x^i$ undergo coordinate transformation as a vector?

Example 1.4.2 $x^i$ is not a vector.
We provide a counter-example that shows that $x^i$ does not satisfy vector equation (1.45). Let $(x,y,z)$ be Cartesian coordinates and $(x',y',z') = (r,\theta,\phi)$ be spherical coordinates. Consider $i' = 3'$. That is, $x^{i'} = x^{3'} = \phi$. In Example 1.1.4 we found that $\phi = \arctan \frac{y}{x}$. So, $\frac{\partial\phi}{\partial x} = \frac{-2\,xy^2}{x^2+y^2}$, $\frac{\partial\phi}{\partial y} = \frac{2\,y}{x^2+y^2}$, $\frac{\partial\phi}{\partial z} = 0$. We wish to show that $\phi = x^{3'} \neq U_j^{3'}\,x^j$ :

$$U_j^{3'}\,x^j = x^1\frac{\partial x^{3'}}{\partial x^1} + x^2\frac{\partial x^{3'}}{\partial x^2} + x^3\frac{\partial x^{3'}}{\partial x^3} = x\frac{\partial\phi}{\partial x} + y\frac{\partial\phi}{\partial y} + z\frac{\partial\phi}{\partial z}$$
$$= \frac{-2\,x^2\,y^2}{x^2+y^2} + \frac{2\,y^2}{x^2+y^2} + 0 = \frac{2\,y^2\,(1-x^2)}{x^2+y^2} \neq \arctan\frac{y}{x} = \phi. \quad ■$$

Theorem 1.4.1 Dot products in Euclidean space are basis-independent.

Proof. $\boldsymbol{\lambda}\cdot\boldsymbol{\lambda} \overset{(1.26)}{=} \lambda^i\,\mu_i$. We wish to show that $\boldsymbol{\lambda}\cdot\boldsymbol{\lambda} = \lambda^{i'}\,\mu_{i'}$.

$$\lambda^{i'}\,\mu_{i'} \overset{(1.45)}{=} U_k^{i'}\,\lambda^k\,U_{i'}^{\ell}\,\mu_\ell = U_{i'}^{\ell}\,U_k^{i'}\,\lambda^k\,\mu_\ell \overset{(1.50)}{=} \delta_k^{\ell}\,\lambda^k\,\mu_\ell = \lambda^\ell\,\mu_\ell = \lambda^i\,\mu_i \quad ■$$

Notation  Because dot products are coordinate-independent, we can let $\underset{'}{G} = (g_{i' j'})$ and $\overset{'}{G} = (g^{i' j'})$ denote the **metric tensor matrices**, where $g_{i' j'} = \mathbf{e}_{i'} \cdot \mathbf{e}_{j'}$ and $g^{i' j'} = \mathbf{e}^{i'} \cdot \mathbf{e}^{j'}$. Compare to $G = (g_{ij})$ and $\hat{G} = (g^{ij})$.

Theorem 1.4.2

$$\boxed{g_{i' j'} = U^k_{i'} U^\ell_{j'} g_{k \ell}}, \quad \text{or} \quad \underset{'}{G} = U^{-1} G\, U \tag{1.52}$$

$$\boxed{g^{i' j'} = U^{i'}_k U^{j'}_\ell g^{k \ell}}, \quad \text{or} \quad \overset{'}{G} = U \hat{G}\, U^{-1} \tag{1.53}$$

Proof. $g_{i' j'} = \mathbf{e}_{i'} \cdot \mathbf{e}_{j'} \overset{(1.4-5)}{=} (U^k_{i'} \mathbf{e}_k) \cdot (U^\ell_{j'} \mathbf{e}_\ell) = U^k_{i'} U^\ell_{j'} \mathbf{e}_k \cdot \mathbf{e}_\ell \overset{(1.22)}{=} U^k_{i'} U^\ell_{j'} g_{k \ell}$ ■

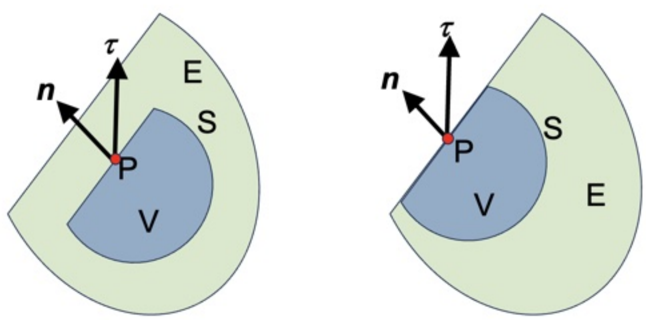
Note  $g_{ij}$ is called a **metric tensor** because it provides access to metric properties such as the lengths of vectors and the angles between them (via the dot product $\boldsymbol{\lambda} \cdot \boldsymbol{\mu} = g_{ij} \lambda^i \mu^j$), and the distance between points (via the line element $ds^2 = g_{ij}\, du^i\, du^j$).

## 1.5 Tensors in Euclidean 3-Space

In Section 1.8 we will define tensors. We will use indices like $a$, $b$, $c$ instead of $i$, $j$, $k$. The definition will include covariant tensors $\tau_a$, $\tau_{a\,b}$, ...; contravariant tensors $\tau^a$, $\tau^{a\,b}$, ...; and mixed tensors $\tau^a_b$, $\tau^{a\,b}_c$, $\tau^a_{b\,c}$, ... . Tensors will be defined as "objects" like $\tau^{a\,b}_{c\ d}$ that transform from an unprimed coordinate system to a primed coordinate system according to pattern definition (1.73).

To motivate the pattern exhibited in this equation, we have already provided pattern equations (1.49) for covariant vectors $\lambda_i$ and contravariant vectors $\lambda^i$, and the metric tensor pattern equations (1.52) and (1.53) for $g_{i\,j}$ and $g^{i\,j}$. In this section we develop equation (1.58), the transformation equation for a mixed tensor $\tau^i_j$. Einstein's field equations are expressed in terms of metric tensors and the stress tensor, so for this pattern definition example we will develop the stress tensor for Euclidean space.

Suppose we have a 3-dimensional elastic body E that is placed under stress by both external and internal forces. It might be helpful to imagine a moon-sized balloon filled with squish-able gel. If the gel is lumpy and massive, then there will be internal gravitational force. If we place the balloon on the Earth, then there will be an external gravitational force as well. If the balloon contains charged particles and it is placed in an electromagnetic field, there will also be internal and external electromagnetic forces. In our development, below, we allow any and all forces.




Let V be a small part of E, let $S$ be the surface of V, and let P be a point of $S$. If P is internal to E, we postulate that the forces on P are all internal. If P is on the surface of E, we postulate that all the forces on P are external.

Forces on P can come from all directions. We can imagine that if the surface *S* were flat, a shear force (i.e., a force that hits *S* at an oblique angle) would have less effect than if the force were perpendicular. Let $\boldsymbol{\tau}$ be the sum of all the pressures (i.e., force per unit area) on *S*. The force at P is a function not only of $\boldsymbol{\tau}$ but also of the unit normal $\boldsymbol{n}$ to S at P. We express this as

$$\mathbf{f} = \boldsymbol{\tau}(\boldsymbol{n}). \tag{1.5-1}$$

We assume $\boldsymbol{\tau}$ to be a linear function so that the total force on V due to stresses can be found by "adding" all the differential forces. That is, total force is defined as

$$\iint_S \boldsymbol{f}\, \mathrm{d}S \equiv \iint_S \boldsymbol{\tau}(\boldsymbol{n})\, \mathrm{d}S. \tag{1.5-2}$$

Since $\boldsymbol{\tau}$ is a linear function, we also have that for vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and scalars $\alpha$ and $\beta$

$$\boldsymbol{\tau}(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) = \alpha\, \boldsymbol{\tau}(\boldsymbol{u}) + \beta\, \boldsymbol{\tau}(\boldsymbol{v}). \tag{1.5-3}$$

Suppose we use curvilinear coordinates $u^i$ to label points of E. Then

$$\boldsymbol{f} = f^i \mathbf{e}_i \tag{1.5-4}$$

and

$$\boldsymbol{n} = n^j \mathbf{e}_j. \tag{1.5-5}$$

Hence,

$$f^i \mathbf{e}_i = n^j\, \boldsymbol{\tau}(\mathbf{e}_j) : \tag{1.54}$$

$$f^i \mathbf{e}_i \overset{(1.5\text{–}4)}{=} \boldsymbol{f} \overset{(1.5\text{–}1)}{=} \boldsymbol{\tau}(\boldsymbol{n}) \overset{(1.5\text{–}5)}{=} \boldsymbol{\tau}(n^j\, \mathbf{e}_j) \overset{(1.5\text{–}3)}{=} n^j\, \boldsymbol{\tau}(\mathbf{e}_j) \quad ✓$$

$\boldsymbol{\tau}(\mathbf{e}_j)$ is a vector so it can be expressed as a linear combination of $\{\mathbf{e}_i\}$:

$$\boldsymbol{\tau}(\mathbf{e}_j) \equiv \tau^i_j\, \mathbf{e}_i : \tag{1.55}$$

Note: we cannot just express $\boldsymbol{\tau}(\mathbf{e}_j)$ as $\boldsymbol{\tau}(\mathbf{e}_j) = \tau^i \mathbf{e}_i$ . This would result in $\boldsymbol{\tau}(\mathbf{e}_1) = \tau^i \mathbf{e}_i$ and also $\boldsymbol{\tau}(\mathbf{e}_2) = \tau^i \mathbf{e}_i$, which would make them equal. Thus, $\tau^i_j$ requires two indices.

As a result, we get

$$f^i = \tau^i_j\, n^j : \tag{1.56}$$

$$\forall\, i,\;\; f^i \mathbf{e}_i \overset{(1.54)}{=} n^j\, \boldsymbol{\tau}(\mathbf{e}_j) \overset{(1.55)}{=} \tau^i_j\, n^j\, \mathbf{e}_i \qquad ✓$$

[The matrix version of this is $\boldsymbol{F} = \boldsymbol{TN}$, where $\boldsymbol{F} = (f^i)$, $\boldsymbol{T} = (\tau^i_j)$, and $\boldsymbol{N} = (n^j)$].

Definition The linear function $\boldsymbol{\tau}$ is called the **stress tensor**. It has components $\tau^i_j$ defined by equation (1.55).

Using primed coordinates $u^{i'}$, we similarly define the primed components $\tau^{i'}_{j'}$ by

$$\boldsymbol{\tau}(\mathbf{e}_{j'}) \equiv \tau^{i'}_{j'}\, \mathbf{e}_{i'}\ ,$$

and that leads to the analogue of equation (1.56):

$$f^{i'} = \tau^{i'}_{j'}\, n^{j'}. \tag{1.57}$$

Because $f^i$ and $n^i$ are vectors, they transform as

$$f^{i'} \overset{(1.49)}{=} U^{i'}_k f^k \tag{1.5-6}$$

and

$$n^{j'} \;=\; U^{j'}_\ell f^\ell\ , \tag{1.5-7}$$

so

$$U^{i'}_k f^k \overset{(1.5\text{–}6)}{=} f^{i'} \overset{(1.57)}{=} \tau^{i'}_{j'}\, n^{j'} \overset{(1.5\text{–}7)}{=} \tau^{i'}_{j'} U^{j'}_\ell n^\ell\ . \tag{1.5-8}$$

Also,

$$U^{i'}_k \tau^k_\ell\, n^\ell \overset{(1.56)}{=} U^{i'}_k f^k \overset{(1.5\text{–}8)}{=} \tau^{i'}_{j'} U^{j'}_\ell n^\ell,$$

and this holds for all unit vectors $\boldsymbol{n}$ at P. So,

$$U^{i'}_k \tau^k_\ell = \tau^{i'}_{j'} U^{j'}_\ell \tag{1.5-9}$$

and, hence,

$$\tau^{i'}_{m'} = U^{i'}_k\ U^\ell_{m'} \tau^k_\ell\ : \tag{1.58}$$

$$U^\ell_{m'} U^{i'}_k \tau^k_\ell \overset{(1.5\text{–}9)}{=} U^\ell_{m'} \tau^{i'}_{j'} U^{j'}_\ell \overset{(1.50)}{=} \tau^{i'}_{j'}\ \delta^{j'}_{m'} = \tau^{i'}_{m'} \qquad ✓$$

Formula (1.58) is the promised transformation formula for the mixed tensor $\tau^i_j$. As required of a tensor formula, repeated indices occur precisely twice, once as a subscript and once as a superscript. The **free indices**, those not involved in summation, carry primes, and the free indices on LHS and RHS match, also as required.

Before delving into curved $N$-manifolds (in Section 1.7), we explore curved surfaces in Euclidean space as examples of curved 2-dimensional manifolds.

## 1.6 Surfaces in Euclidean 3-space

Just as a curve $\gamma$ in Euclidean 3-space can be defined using a single parameter $t$, a surface $\Sigma$ can be defined using a pair of parameters $u$ and $v$. We will assume that $\Sigma$ is differentiable at every point. $\Sigma$ can be described parametrically

$$x = x(u,v), \;\; y = y(u,v), \;\; z = z(u,v) \tag{1.59}$$

or by a single equation

$$\boldsymbol{r} = x(u,v)\,\boldsymbol{i} + y(u,v)\,\boldsymbol{j} + z(u,v)\,\boldsymbol{k}\,.$$

At each point P = $(u_0, v_0) \in \Sigma$ there are two level curves in $\Sigma$, defined 'parametrically'

$$u = u_0 \;\text{ and }\; v = v_0$$

or by single equations

$$\boldsymbol{r}(u) = x(u,v_0)\,\boldsymbol{i} + y(u,v_0)\,\boldsymbol{j} + z(u,v_0)\,\boldsymbol{k} \quad\text{and}\quad \boldsymbol{r}(v) = x(u_0,v)\,\boldsymbol{i} + y(u_0,v)\,\boldsymbol{j} + z(u_0,v)\,\boldsymbol{k}\,.$$

Before we discuss bases, we need to discuss vectors. We are only concerned with vectors that point in the directions of possible movement in $\Sigma$. If $\Sigma$ is curved at a point P, vectors emanating from P do not lie in $\Sigma$. Movement in $\Sigma$ away from P is best described by the vectors that lie in the tangent plane to $\Sigma$ at P. Consequently, the only vectors that are considered when analyzing surfaces are those that lie in the tangent planes.

Definition A **field** is an assignment of a value to every point in a space. If the value is a scalar, then it is a **scalar field**. If the value is a vector or a tensor, then the field is a **vector field** or **tensor field**, respectively. For a **vector field on a surface**, we also require that each vector be tangential to the surface.

Definition The **natural basis** of a vector field at a point P on a surface $\Sigma$ is defined as

$$\boldsymbol{e}_u \equiv \frac{\partial \boldsymbol{r}}{\partial u} \;\text{ and }\; \boldsymbol{e}_v \equiv \frac{\partial \boldsymbol{r}}{\partial v}.$$

Suppose $\boldsymbol{\lambda}$ is an element of a vector field on $\Sigma$. We can express $\boldsymbol{\lambda}$ as a linear combination of the natural basis vectors:

$$\boldsymbol{\lambda} = \lambda^u\,\boldsymbol{e}_u + \lambda^v\,\boldsymbol{e}_v \tag{1.60}$$

$\boldsymbol{e}_u$ is tangent to the curve $v = v_0$ shown in Figure 1.4 because $\boldsymbol{e}_u = \dfrac{\partial \boldsymbol{r}}{\partial u}$ is the derivative along the curve $r(u,v_0)$ of constant $v_0$. Similarly, $\boldsymbol{e}_v$ is tangent to the curve $u = u_0$ . Both are tangential to the surface $\Sigma$ and together they define a tangent plane to $\Sigma$ at P.

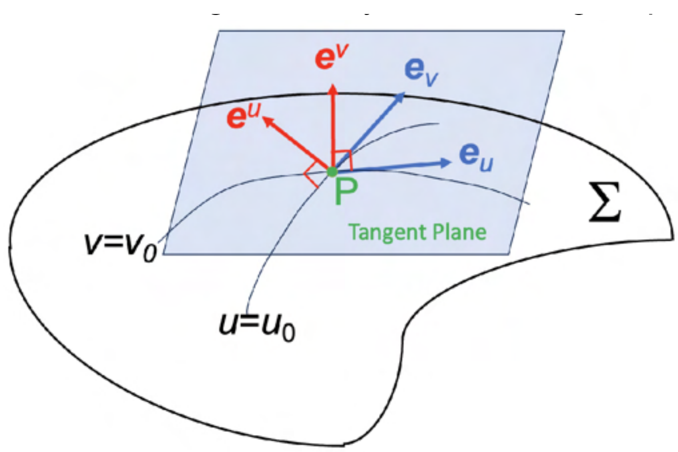



Figure 1.4 Tangent plane to Σ at P

There is also a **dual basis** $\{\mathbf{e}^u, \mathbf{e}^v\}$ for the surface Σ but the natural instinct to define it as $(\nabla u, \nabla v)$ using the gradient formula (1.9) doesn't work because those formulas may not have meaning:

> In (1.9) we are given $x = x(u,v)$, $y = y(u,v)$, and $z = z(u,v)$. For a surface, it is not always possible to solve for $u = u(x,y,z)$ and $v = v(x,y,z)$ even in principle. This is because we would be solving 3 equations in the 2 unknowns $u$ and $v$, and it is possible that the extra equation is inconsistent with the first two.

This inconsistency will be illustrated in Example 1.6.1. Since we cannot solve for $u$ and $v$, then the formulas (1.9) for $\nabla u$ and $\nabla v$ have no meaning for surfaces.

Nonetheless, $u$ and $v$ are scalar fields on Σ. Gradients, meaning direction of steepest approach, always exists for scalar fields. We simply need another way solve for them. Our approach is to define the vectors $\nabla u$ and $\nabla v$ by specifying their directions and magnitudes.

As previously pointed out, $\mathbf{e}_v$ is tangent to the flat curve $u = u_0$. So, $\mathbf{e}_v$ points in the direction of 'no change in $u$'. The direction of $\mathbf{e}^u$, the direction of steepest approach, is normal to $\mathbf{e}_v$. We choose the normal that points in the direction of increasing $u$ as shown in Figure 1.4. Similarly the direction of $\mathbf{e}^v$ is the normal to $\mathbf{e}_u$ that points in the direction of increasing $v$. Normality is expressed as

$$\mathbf{e}^u \cdot \mathbf{e}_v = 0 \text{ and } \mathbf{e}^v \cdot \mathbf{e}_u = 0. \tag{1.61}$$

Remember, there are only 2 normals to consider in each case because we only permit vectors that lie in the tangent plane to Σ at P. By choosing the direction of increasing $u$ (or $v$), we identify a unique direction for each basis vector.

We specify the magnitudes of $\mathbf{e}^u$ and $\mathbf{e}^v$ by requiring

$$\mathbf{e}^u \bullet \mathbf{e}_u = 1, \quad \mathbf{e}^v \bullet \mathbf{e}_v = 1. \tag{1.62}$$

Since our index convention for *i*, *j*, and *k* is *i*, *j*, *k* = 1 - 3, we will use the convention *A, B, C* = 1 -2 for surfaces. Then, the natural basis $\{\mathbf{e}_u, \mathbf{e}_v\}$ can be denoted $\{\mathbf{e}_A\}$, the dual basis $\{\mathbf{e}^u, \mathbf{e}^v\}$ can be denoted $\{\mathbf{e}^A\}$, and equations (1.61) and (1.62) can be consolidated as

$$\mathbf{e}^A \bullet \mathbf{e}_B = \delta^A_B, \tag{1.63}$$

which is the analog of equation (1.17).

If $\boldsymbol{\lambda}$ is a vector defined on a surface Σ, then it can be expressed as

$$\boldsymbol{\lambda} = \lambda^A \mathbf{e}_A = \lambda_A \mathbf{e}^A \tag{1.6-1}$$

where

$$\lambda^A = \boldsymbol{\lambda} \bullet \mathbf{e}^A \text{ and } \lambda_A = \boldsymbol{\lambda} \bullet \mathbf{e}_A, \tag{1.6-2}$$

which are the analogs of equations (1.16) and (1.19 – 1.20). Remember, we only consider vectors $\boldsymbol{\lambda}$ that lie in the tangent plane at P.

Similar to definitions 1.22 and 1.24, we define the metric tensors

$$g_{AB} = \mathbf{e}_A \bullet \mathbf{e}_B \text{ and } g^{AB} = \mathbf{e}^A \bullet \mathbf{e}^B. \tag{1.6-3}$$

Continuing to mimic the development in Section 1.2, we also get

$$\boldsymbol{\lambda} \bullet \boldsymbol{\mu} \overset{(1.26)}{=} \lambda_A \mu^A = \lambda^A \mu_A = g_{AB} \lambda^A \mu^B = g^{AB} \lambda_A \mu_B, \tag{1.6-4}$$

$$g^{AB} \lambda_B \overset{(1.27)}{=} \lambda^A \text{ and } g_{AB} \lambda^B \overset{1.28}{=} \lambda_A, \tag{1.6-5}$$

$$\mathbf{e}^A \overset{(1.2-3)}{=} g^{AB} \mathbf{e}_B \text{ and } \mathbf{e}_A \overset{(1.2-4)}{=} g_{AB} \mathbf{e}^B, \tag{1.6-6}$$

$$g^{AB} g_{BC} \overset{(1.29)}{=} \delta^A_C \text{ and } g_{AB} g^{BC} \overset{(1.31)}{=} \delta^C_A. \tag{1.6-7}$$

Since we have an explicit formula for the natural basis but not for the dual basis, in practice how do we compute the dual basis? The next example shows that we can first compute the natural basis, then generate the covariant metric tensor from the natural basis, next invert the tensor to generate the contravariant metric tensor, and finally use that tensor to compute the dual basis.

Example 1.6.1 Compute the dual basis for the hyperbolic paraboloid surface Σ defined by

$$\boldsymbol{r} = (u + v)\,\boldsymbol{i} + (u - v)\,\boldsymbol{j} + 2\,u\,v\,\boldsymbol{k},$$

where $-\infty < u, v < \infty$. Show that 3-space equations (1.9) for $\nabla u$ and $\nabla v$ are not valid for this surface, illustrating that another method must be found to define the dual basis.

The single equation for Σ is $\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$, which generates the parametric equations

$$x = u + v \qquad y = u - v \qquad z = 2uv$$
$$u = \tfrac{1}{2}(x + y) \qquad v = \tfrac{1}{2}(x - y) \qquad uv = \tfrac{1}{2}z\,.$$

We see that that $\frac{\partial u}{\partial z} = 0 = \frac{\partial v}{\partial z}$, so $z = 2uv$ yields

$1 = \frac{\partial z}{\partial z} = \frac{\partial(2\,u\,v)}{\partial z} = 2\,(u\frac{\partial v}{\partial z} + v\frac{\partial u}{\partial z}) = 2\,(0 + 0) = 0$ # The equations for $u$ and $v$ are inconsistent and cannot be solved. ✔

The natural basis is

$$\boldsymbol{e}_u \equiv \frac{\partial \boldsymbol{r}}{\partial u} \overset{1.2-1}{=} \frac{\partial x}{\partial u}\,\boldsymbol{i} + \frac{\partial y}{\partial u}\,\boldsymbol{j} + \frac{\partial z}{\partial u}\,\boldsymbol{k} = \boldsymbol{i} + \boldsymbol{j} + 2\,v\,\boldsymbol{k}$$
$$\boldsymbol{e}_V \equiv \frac{\partial \boldsymbol{r}}{\partial v} \overset{1.2-1}{=} \frac{\partial x}{\partial v}\,\boldsymbol{i} + \frac{\partial y}{\partial v}\,\boldsymbol{j} + \frac{\partial z}{\partial v}\,\boldsymbol{k} = \boldsymbol{i} - \boldsymbol{j} + 2\,u\,\boldsymbol{k}.$$

So, the covariant metric tensor matrix is

$$(g_{AB}) \overset{1.22}{=} \begin{pmatrix} \boldsymbol{e}_u \bullet e_u & \boldsymbol{e}_u \bullet e_v \\ \boldsymbol{e}_v \bullet e_u & \boldsymbol{e}_v \bullet e_v \end{pmatrix} = 2\begin{pmatrix} 1+2\,v^2 & 2\,u\,v \\ 2\,u\,v & 1+2\,u^2 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}.$$

From (1.6-7), the contravariant metric tensor matrix is

$$(g^{AB}) = (g_{AB})^{-1} = \frac{1}{2\left(1+2\,u^2+2\,v^2\right)}\begin{pmatrix} 1+2\,u^2 & -2\,u\,v \\ -2\,u\,v & 1+2\,v^2 \end{pmatrix} = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix}.$$

Finally,

$$\boldsymbol{e}^u = \boldsymbol{e}^1 = g^{1A}\,\boldsymbol{e}_A = g^{11}\,\boldsymbol{e}_1 + g^{12}\,\boldsymbol{e}_2 = g^{11}\,\boldsymbol{e}_u + g^{12}\,\boldsymbol{e}_v$$
$$= \frac{1}{2\left(1+2\,u^2+2\,v^2\right)}\left[(1 + 2\,u^2 - 2\,u\,v)\ \boldsymbol{i} + (1 + 2\,u^2 + 2\,u\,v\,)\,\boldsymbol{j} + 2v\,\boldsymbol{k}\right]$$

$$\boldsymbol{e}^v = \frac{1}{2\left(1+2u^2+2v^2\right)}\left[\left(1+2v^2-2uv\right)\boldsymbol{i} - \left(1+2v^2+2uv\right)\boldsymbol{j} + 2u\,\boldsymbol{k}\right]$$

Check: $$\boldsymbol{e}^u \bullet \boldsymbol{e}_u = \frac{\left(1+2u^2-2uv\right)+\left(1+2u^2+2uv\right)+4v^2}{2\left(1+2u^2+2v^2\right)} = 1 \quad ✓$$

$$\boldsymbol{e}^v \bullet \boldsymbol{e}_v = \frac{\left(1+2v^2-2uv\right)+\left(1+2v^2+2uv\right)+4u^2}{2\left(1+2u^2+2v^2\right)} = 1 \quad ✓ \qquad \blacksquare$$

Curves in Σ are developed similarly to curves in space. If $u$ and $v$ are differential functions of $t$ on some interval I = [$a$,$b$], then $x = x(\,u(t),\, v(t)\,) = x(t)$, $y = y(t)$, and $z = z(t)$ are also differential functions of $t$ on I. Thus,

$$\boldsymbol{r} = x(t)\,\boldsymbol{i} + y(t)\,\boldsymbol{j} + z(t)\,\boldsymbol{k} = x(u,v)\,\boldsymbol{i} + y(u,v)\,\boldsymbol{j} + z(u,v)\,\boldsymbol{k}$$

traces out a curve $\gamma$, and $\frac{d\boldsymbol{r}}{dt}$ is tangent to $\gamma$ at each point ($u$, $v$).

$$\dot{\boldsymbol{r}}(t) = \frac{d\boldsymbol{r}}{dt} = \frac{\partial \boldsymbol{r}}{\partial u}\frac{du}{dt} + \frac{\partial \boldsymbol{r}}{\partial v}\frac{dv}{dt} = \frac{du}{dt}\,\boldsymbol{e}_u + \frac{dv}{dt}\,\boldsymbol{e}_v$$

$$d\boldsymbol{r} = \dot{\boldsymbol{r}}(t)\,dt = du\,\boldsymbol{e}_u + dv\,\boldsymbol{e}_v$$

$$\boxed{d\boldsymbol{r} = du^A\,\boldsymbol{e}_A} \qquad (1.6\text{-}8)$$

$$\boxed{ds^2 = d\boldsymbol{r} \bullet d\boldsymbol{r} = g_{AB}\,\dot{u}^A\,\dot{u}^B\,d\,t^2}: \qquad (1.65)$$

$$ds^2 \overset{(1.39)}{=} d\boldsymbol{r} \bullet d\boldsymbol{r} = du^A\,\boldsymbol{e}_A \bullet du^B\,\boldsymbol{e}_B \overset{(1.6-3)}{=} g_{AB}\,du^A\,du^B = g_{AB}\,\dot{u}^A\,\dot{u}^B\,dt^2 \quad ✓$$

As $t$ increases from $a$ to $b$, $ds > 0$. Thus, $ds = +\sqrt{ds^2}$, and the length of the curve is

$$\boxed{L = \int_\gamma ds = \int_a^b \sqrt{g_{AB}\,\dot{u}^A\,\dot{u}^B}\,dt}. \qquad (1.64)$$

Later sections will use the results of Exercise 1.6.2, stated below.

Exercise 1.6.2 Develop the equations of line elements.

| | |
|---|---|
| (a) Sphere | $ds^2 = r^2\,(d\theta^2 + \sin^2\theta\,d\phi^2)$ |
| (b) Cylinder | $ds^2 = r^2\,d\phi^2 + dz^2$ |
| (c) Hyperbolic paraboloid | $ds^2 = (2 + 4\,v^2)\,du^2 + (2 + 4\,u^2)\,dv^2 + 8uv\,du\,dv$ |

Example 1.6.2 (not in book) Show that the Euclidean space line element in spherical coordinates is

$$ds^2 \overset{(1.83)}{=} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta \, d\phi^2. \qquad (1.6\text{-}9)$$

Solution.

In flat space, $\boldsymbol{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, where in spherical coordinates

$x = r\sin\theta\cos\phi,\;\; y = r\sin\theta\sin\phi,\;\; z = r\cos\theta$

From Example 1.1.4, $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$, and

$$\mathbf{e}_1 = \mathbf{e}_{\boldsymbol{r}} = \frac{\partial \boldsymbol{r}}{\partial r} = \frac{\partial x}{\partial r}\boldsymbol{i} + \frac{\partial y}{\partial r}\boldsymbol{j} + \frac{\partial z}{\partial r}\boldsymbol{k} = \sin\theta\cos\phi\,\boldsymbol{i} + \sin\theta\sin\phi\,\boldsymbol{j} + \cos\theta\,\boldsymbol{k}$$

$$\mathbf{e}_2 = \mathbf{e}_\theta = \frac{\partial \boldsymbol{r}}{\partial \theta} = \frac{\partial x}{\partial \theta}\boldsymbol{i} + \frac{\partial y}{\partial \theta}\boldsymbol{j} + \frac{\partial z}{\partial \theta}\boldsymbol{k} = r\cos\theta\cos\phi\,\boldsymbol{i} + r\cos\theta\sin\phi\,\boldsymbol{j} - r\sin\theta\,\boldsymbol{k}$$

$$\mathbf{e}_3 = \mathbf{e}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = \frac{\partial x}{\partial \phi}\boldsymbol{i} + \frac{\partial y}{\partial \phi}\boldsymbol{j} + \frac{\partial z}{\partial \phi}\boldsymbol{k} = -r\sin\theta\sin\phi\,\boldsymbol{i} + r\sin\theta\cos\phi\,\boldsymbol{j}$$

$g_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j \;\Rightarrow\; g_{11} = 1$, $g_{22} = r^2$, $g_{33} = r^2 \sin^2\theta$, and all other $g_{ij} = 0$.

Therefore, the Euclidean 3-space line element is

$$ds^2 \overset{(1.83)}{=} g_{ij}\, dx^i\, dx^j = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2. \qquad \blacksquare$$

## 1.7 Manifolds

We seek an environment in which we can define coordinate systems and take partial derivatives of one coordinate system in terms of another. We also need to be able to discuss orthogonality. The structure we seek is a differentiable manifold with a metric tensor field. (Having a metric tensor field is not a property of a general manifold.)
A metric tensor provides the concept of distance. It also generates an inner product, which provides the concept of angles. Metric tensors will be defined in Section 1.8. In this section we define manifolds and differential manifolds.

Conventions We already have conventions that Euclidean 3-space indices are *i*, *j*, *k*, ... and range over 1, 2, and 3, and that upper case indices *A*, *B*, *C*, ... are used for 2-dimensional surfaces, and range over 1, 2. We now extend these conventions to include *N*-dimensional manifolds (below) and spacetime (Appendix A).

| **Object** | **Indices** | **Range** | **Coordinates** | **Jacobian Matrices** |
|---|---|---|---|---|
| Euclidean 3-space | $i, j, k, \ldots$ | 1, 2, 3 | $(x^i)$, $(u^i)$ | $U^{i'}_{j}$ |
| Surface | $A, B, C, \ldots$ | 1, 2 | $(u^A)$ | – |
| Manifold | $a, b, c, \ldots$ | $1, 2, 3, \ldots, N$ | $(x^a)$ | $X^{a'}_{b}$ |
| Spacetime | $\mu, \nu, \sigma, \ldots$ | 0, 1, 2, 3 | $(x^\mu)$ | $X^{\mu'}_{\nu}$ |

Definition $\mathbb{R}^N$ is defined to be the set of points $\boldsymbol{x} = (x^a) = \left(x^1, x^2, \ldots, x^N\right)$, where $x^a \in \mathbb{R}$.

$\mathbb{R}^N$ has a natural topology (collection of open set) that is generated from the open rectangular *N*-cubes $(d^1 - c^1) \times (d^2 - c^2) \times \ldots \times (d^N - c^N)$, where $c^a$, $d^a \in \mathbb{R}$. The rules for generating all the open sets from the open *N*-cubes are that any union of open sets is open and all finite intersections of open sets are open.

The more standard method is is not to use *N*-cubes but to generate the open sets by starting with the open *N*-balls. However, that approach requires a metric, because an *N*-ball is defined as the set of points a fixed distance (a metric) from a center point. One class of manifolds we are about to define will have a metric, but the other class will not, so we must defer referring to metrics for now.

We will use $\mathbb{R}^N$ both to denote the set of points and the topological space (points plus open sets). It should be clear from context which we are discussing. When we finally do define the Euclidean metric for $\mathbb{R}^N$ in Section 1.9, we will denote the resulting metric space Euclidean $N$-space, $\mathbb{E}^N$ to distinguish it from the more general topological space $\mathbb{R}^N$ we use now.

Definition An ***N*-dimensional manifold** is an $N$-dimensional topological space M that is locally homeomorphic to the topological space $\mathbb{R}^N$.

"Locally" simply means that each point in M is contained in an open set that is homeomorphic to $\mathbb{R}^N$. Homeomorphic means that there is a 1-1 function from M to $\mathbb{R}^N$ that maps open sets to open sets. This establishes the structure necessary to discuss and develop continuous and differentiable functions. (A continuous function $f$ is defined as one whose inverse, $f^{-1}$, maps open sets to open sets.)

Definition A **differentiable manifold**, M, is a manifold in which the open neighborhoods U have differentiable coordinate systems. That is, if $(x^1, \ldots, x^N)$ is a coordinate system on U and $(x^{1'}, \ldots, x^{N'})$ is a coordinate system on an overlapping open neighborhood U′ then on U∩U′ we can express

$$x^{a'} = x^{a'}(x^1, \ldots, x^N), \quad (a' = 1, \ldots, N) \tag{1.66}$$

$$x^a = x^a\left(x^{1'}, \ldots, x^{N'}\right), \quad (a = 1, \ldots, N) \tag{1.67}$$

where the functions $x^a$ and $x^{a'}$ are differentiable, so that the **Jacobian matrix** and its inverse both exist:

$$X^{a'}_b \equiv \frac{\partial x^{a'}}{\partial x^b} \tag{1.7-1}$$

$$X^a_{b'} \equiv \frac{\partial x^a}{\partial x^{b'}}. \tag{1.7-2}$$

Convention We will not only assume that all manifolds are differentiable, but analytic (infinitely differentiable). In particular, the order of differentiation does not matter.

Definitions $X = (X^{a'}_b)$ is the **Jacobian matrix** associated with equations (1.67) and it provides the change of coordinates from $x^a$ to $x^{a'}$. The transpose is its inverse: $X^{-1} = X^{\mathrm{T}} = (X^b_{a'})$. The **Jacobian determinants** are det X and det $X^{-1}$.

The chain rule $\frac{\partial x^a}{\partial x^{b'}}\frac{\partial x^{b'}}{\partial x^c} = \frac{dx^a}{dx^c} = \delta^a_c$ justifies that the transpose is the inverse:

$$\boxed{X^a_{b'} X^{b'}_c = \delta^a_c} \tag{1.68}$$

$$\boxed{X^{a'}_b X^b_{c'} = \delta^{a'}_{c'} = \delta^a_c}\,. \tag{1.69}$$

$\delta^{a'}_{c'} = \delta^a_c$ because the indices of the delta function are dummy indices. Since X is invertible in $U \cap U'$, det $X \neq 0$ in $U \cap U'$.

We have just shown that if $(x^a)$ and $\left(x^{a'}\right)$ are coordinate systems in regions U and U′ around a point P, then there is an overlap region where the Jacobian det $(X^{a'}_b)$ is non-zero. There is a partial converse that gives the conditions for a set of expressions $\left(x^{a'}\right)$ to be a coordinate system.

Theorem 1.7.1 Suppose $(x^a)$ is a coordinate system in a neighborhood $U$ about a point P, and $\left(x^{a'}\right)$ is a set of continuously differentiable equations in $U$. Then $\left(x^{a'}\right)$ is a coordinate system iff there is a neighborhood $U'$ about P where det $(X^{a'}_b) \neq 0$.

Proof outline for the converse. The Inverse Function Theorem posits that P has a neighborhood U′ such that the function $x^a \mapsto x^{a'}$ is 1-1. Since the coordinates $(x^a)$ are in 1-1 correspondence with the points of U′, then the expressions $\left(x^{a'}\right)$ are also in 1-1 correspondence with the points of U′. That is, $\left(x^{a'}\right)$ is a coordinate system on U′. ■

In Section 1.4 we defined vectors in terms of the natural and dual bases and then showed that equation (1.45) holds. Basis-free definitions are preferred in general relativity, so in this section we do the opposite: we take coordinate transformation equation (1.45), now updated to manifold notation and renamed (1.70), below, as the defining condition for an object to be a vector, and we develop the bases later, as an afterthought.

This approach is convoluted and, at times, not fully explained in the book in order to avoid being side-tracked. However, it allows the authors to focus on tensor coefficients like $T^{ab}_{cd}$ and how they are manipulated, and to ignore the bases. The coefficients reflect the tensor magnitude, and the bases reflect the tensor direction. However, tensor direction can be obtained, also, from coefficients by using metric tensor and arc-length, a discussion we postpone until Section 1.9 (see equation 1.81).

Convention  We will often refer to an object $\boldsymbol{\lambda}$ by its components, $\lambda_a$.

Definition  Let P be a point with a neighborhood U that has overlapping unprimed and primed coordinate systems. A **contravariant vector** at a point P is an object $\boldsymbol{\lambda}$ such that if $\boldsymbol{\lambda}$ has components $\lambda^a$ in an unprimed coordinate system then $\boldsymbol{\lambda}$ has components $\lambda^{a'}$ in the primed coordinate system, where the components transform as

$$\lambda^{a'} = X^{a'}_b \lambda^b. \tag{1.70}$$

A **covariant vector** is an object $\boldsymbol{\lambda}$ such that if $\boldsymbol{\lambda}$ has components $\lambda_a$ in the unprimed coordinate system then $\boldsymbol{\lambda}$ has components $\lambda_{a'}$ in the primed coordinate system, where

$$\lambda_{a'} = X^b_{a'} \lambda_b. \tag{1.71}$$

A **vector field** is a set of vectors that are defined at each point of the manifold M.

Theorem 1.7.2  Suppose P = $(x^a) \in$ M is a point in an open set U having overlapping unprimed and primed coordinate systems. Let $\gamma$ be a curve through P described parametrically by differentiable functions $x^a = x^a(t)$ for $t$ in some interval $\mathbf{I}$. Then the object $\dot{\boldsymbol{x}}$ having components $\dot{x}^a(t)$ is a contravariant vector.

Proof.  In a primed coordinate system, $\gamma$ is given by

$$x^{a'}(t) \overset{1.66}{=} x^{a'}(x^1, ..., x^N). \tag{1.72}$$

By the Chain Rule,

$$\dot{x}^{a'} \equiv \frac{dx^{a'}}{dt} = \frac{\partial x^{a'}}{\partial x^b} \frac{dx^b}{dt} \overset{1.7-1}{=} X^{a'}_b \dot{x}^b.$$

This shows that $\dot{x}^a$ transforms according to equation (1.70), which means that $\dot{x}^a(t)$ is a contravariant vector. ■

We loosely refer to $\dot{x}^a$ as  the the vector $\dot{\boldsymbol{x}}$.

Definition  The contravariant vector $\dot{x}^a$ is the **tangent vector to $\boldsymbol{\gamma}$**. (1.7-3)

Definition  A **scalar** is a real-valued function defined at a point P. Given a coordinate system $(x^a)$, a scalar is expressed as a function $\varphi : (x^a) \to \mathbb{R}$. A **scalar field** is a set of scalars that are defined at each point of the manifold M. (1.7-4)

Theorem 1.7.3  Let $\varphi = \varphi(x^a)$ be a differentiable scalar field on M . Define $\partial_a \varphi \equiv \dfrac{\partial \varphi}{\partial x^a}$. Then $\partial_a \varphi$ is a covariant vector.

Proof.  In a primed coordinate system,

$$\partial_{a'} \varphi \equiv \frac{\partial \varphi}{\partial x^{a'}} = \frac{\partial \varphi}{\partial x^b} \frac{\partial x^b}{\partial x^{a'}} \overset{1.7-2}{=} X^b_{a'} \partial_b \varphi.$$

This shows that $\partial_a \varphi$ transforms according to equation (1.71), which means that $\partial_a \varphi$ is a covariant vector. ■

Definition  The covariant vector $\partial_a \varphi$ is the **gradient of $\boldsymbol{\phi}$**. (1.7-5)

In Euclidean 3-space, $\nabla \phi = \partial_x \phi \, \boldsymbol{i} + \partial_y \varphi \, \boldsymbol{j} + \partial_z \varphi \, \boldsymbol{k}$ and, thus, has components $\partial_i \varphi$, consistent with the gradient definition (1.7-5) for manifolds.

In manifold M, the function $\varphi$ that picks out the $b$th coordinate of each point of M is an example of a scalar field. That is, if $\varphi(x^a) \equiv x^b$, then then vector $\partial_a \varphi$ is the gradient of $\varphi$.

## 1.8 Tensor Fields on Manifolds

We define tensors as generalized vectors in the sense that they are defined as objects $\tau^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ that obey equation (1.73), below, a generalized form of equations (1.70–1.71) that were used to define vectors. Since the coefficients of tensors, like vectors, carry most of the information, the bases are largely ignored in practice. However, for the sake of completeness, they will be developed after-the-fact in Section 1.10.

Definition  Let P = $(x^1, ..., x^N)$ be a point in M having a neighborhood U that possesses both an unprimed and a primed coordinate system. A **type (*r,s*) tensor** at the point P is an object $\boldsymbol{\tau}$ that consists of real components $\tau^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ that transform to $\boldsymbol{\tau}'$ having components $\tau^{a_1{}' \ldots a_r{}'}_{b_1{}' \ldots b_s{}'}$ according to

$$\tau^{a_1{}' \ldots a_r{}'}_{b_1{}' \ldots b_s{}'} = X^{a_1{}'}_{c_1} \cdots X^{a_r{}'}_{c_r} \, X^{d_1}_{b_1{}'} \cdots X^{d_s}_{b_s{}'} \, \tau^{c_1 \ldots c_r}_{d_1 \ldots d_s}, \tag{1.73}$$

where the Jacobian matrices $X^{a_k{}'}_{c_k} = \left(\frac{\partial x^{a_k{}'}}{\partial x^{c_k}}\right)$ and $X^{d_l}_{b_l{}'} = \left(\frac{\partial x^{d_l}}{\partial x^{l'}}\right)$ are evaluated at the point P. We will loosely denote a tensor as $\boldsymbol{\tau} = \tau^{a_1 \ldots a_r}_{b_1 \ldots b_s}$. The **rank** of the tensor is (*r*,*s*). If $s = 0$ the tensor is a **contravariant tensor.** If $r = 0$ the tensor is a **covariant tensor.** If $r, s \neq 0$ then the tensor is called **mixed.** A type (0,0) tensor is a **scalar**, defined as a frame-independent function defined on the coordinates:

$$\varphi = \varphi(x^a) \equiv \varphi(x^1, \cdots, x^n) = \varphi(x^{a'}), \text{ where } (x^{a'}) \text{ is an alternate coordinate system.}$$

If a type (*r*,*s*) tensor is defined at every point of M, the result is a **tensor field**. A **differentiable tensor field** is one in which the tensors are differentiable with respect to the coordinates $x^a$.

Examples of equation (1.73) for low rank (*r*,*s*) tensors

**(0,0)**: $\boldsymbol{\tau} = \tau$, a scalar, a frame-independent function that maps coordinates to scalars.

**(1,0)**: Rewrite vector equation (1.70) as $\tau^{a'} = X^{a'}_c \tau^c$, a contravariant vector (1.8-1)

**(0,1)**: Rewrite vector equation (1.71) as $\tau_{b'} = X^d_{b'} \tau_d$, a covariant vector (1.8-2)

**(2,0)**: $\tau^{a' b'} = X^{a'}_c X^{b'}_d \tau^{c d}$, a contravariant tensor (1.74)

**(0,2)**: $\tau_{a' b'} = X^c_{a'} X^d_{b'} \tau_{c d}$, a covariant tensor (1.8-3)

**(1,1)**: $\tau^{a'}_{b'} = X^{a'}_c X^d_{b'} \tau^c_d$, a mixed tensor (1.8-4)

**(1,2)**: $\tau^{a'}_{b' c'} = X^{a'}_d X^e_{b'} X^f_{c'} \tau^d_{e f}$, another mixed tensor (1.75)

If a tensor is defined at every point P, and M is a curve $\gamma(t)$, the result is a **tensor field along $\boldsymbol{\gamma}$** and we can regard the tensor components as functions of *t*. If M is a surface Σ, the result is a **tensor field on a surface** and we can regard the tensor components as functions of *u* and *v.*

There are four operations on tensors.

1. **Addition**: Addition of tensor components of the same type yields a tensor of the same type
2. **Scalar multiplication:** Multiplying a tensor by a scalar yields a tensor of the same type
3. **Tensor product:** The tensor product of two tensors is a tensor. It is an outer product, forming all possible product combinations. Tensor products are denoted by juxtaposition (see example, below).
4. **Contraction** is the cancelation of tensor indices due to summation. Contraction of a tensor yields a tensor of lesser rank.

Example Tensor Product Suppose $\lambda_b$ and $\tau^a_c$ are the components of type (0,1) and type (1,1) tensors $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$, respectively. Form a type (1,2) object $\boldsymbol{\sigma}$ having components $\sigma^a_{\ b\,c} \equiv \lambda_b\, \tau^a_c$ . Then $\boldsymbol{\sigma}$ is a tensor:

$$\lambda_{b'} \overset{1.71}{=} X^e_{b'}\, \lambda_e \quad \text{and} \quad \tau^{a'}_{c'} \overset{1.8\text{–}4}{=} X^{a'}_d\, X^f_{c'}\, \tau^d_f\, .$$

$$\sigma^{a'}_{\ b'\,c'} = \lambda_{b'}\, \tau^{a'}_{c'} = X^e_{b'}\, \lambda_e\, X^{a'}_d\, X^f_{c'}\, \tau^d_f = X^{a'}_d\, X^e_{b'}\, X^f_{c'}\, \sigma^d_{\ e\,f}\, .$$

This shows that the tensor product, $\sigma^a_{\ b\,c}$ , satisfies equation (1.75) .
Hence, $\boldsymbol{\sigma}$ is a type (1,2) tensor. ✓

The phrase “outer product” in the above definition of tensor product means that $\boldsymbol{\sigma}$ possesses components formed from every combination of *a*, *b*, *c* = 1, 2, ... , *n*.

Example Contraction in a tensor: Suppose $\tau^a_{\ b\,c}$ are the components of a type (1,2) tensor $\boldsymbol{\tau}$, and we use contraction to form $\boldsymbol{\sigma}$, a type (0,1) object having components $\sigma_b \equiv \tau^a_{\ b\,a}$ . Then $\boldsymbol{\sigma}$ is a (0,1) tensor:

$$\sigma_{b'} = \tau^{a'}_{\ b'\,a'} \overset{(1.75)}{=} X^{a'}_d\, X^e_{b'}\, X^f_{a'}\, \tau^d_{\ e\,f} = \delta^f_d\, X^e_{b'}\, \tau^d_{\ e\,f} = X^e_{b'}\, \tau^d_{\ e\,d} = X^e_{b'}\, \sigma_e\, ,$$

which transforms $\sigma_b$ as a (0,1) tensor by (1.8-2).

Example Contraction in a tensor product: Define $\sigma^a_{\ c\,d} \equiv \lambda^{a\,b}_{c\,d}\, \tau_b$. This represents the contraction of a rank (2,3) tensor, $\lambda^{a\,b}_{c\,d}\, \tau_e$, into a rank (1,2) tensor, $\sigma^a_{\ c\,d}$.

The next theorem roughly states that if a "tensor candidate" generates a tensor when contraction "kills off" indices, then the tensor candidate is an actual tensor. We provide a precise statement and proof of this theorem for a (1,2) tensor because the general case has many messy indices. This version of the theorem serves to illustrate the general theorem and its proof. The full theorem holds even if $\tau$ and $\lambda$ have fewer or more indices and even if there are additional tensor terms.

Quotient Theorem 1.8.1 Suppose that at a point P there are scalars $\tau^a_{b\,c}$ such that for any tensor component $\lambda^a$, the contraction $\tau^a_{b\,c}\,\lambda^c$ is a tensor. Then $\boldsymbol{\tau} = \tau^a_{b\,c}$ is a tensor.

Proof. Let $\lambda^a$ be an arbitrary vector component. By hypothesis, $\sigma^a_b \equiv \tau^a_{b\,c}\lambda^c$ is a tensor. That is, it transforms according to equation (1.8-4) as

$$\sigma^{a'}_{b'} = X^{a'}_d\, X^e_{b'}\, \sigma^d_e.$$

Replacing $\sigma$ on both sides of the equation yields

$$\tau^{a'}_{b'\,c'}\,\lambda^{c'} = X^{a'}_d\, X^e_{b'}\, \tau^d_{e\,f}\,\lambda^f. \tag{1.76}$$

The vector $\lambda^a$ transforms as $\lambda^{c'} = X^{c'}_f\,\lambda^f$. Plugging this into equation (1.76) gives

$$\tau^{a'}_{b'\,c'}\,X^{c'}_f\,\lambda^f = X^{a'}_d\, X^e_{b'}\, \tau^d_{e\,f}\,\lambda^f.$$

Since this equation holds for any vector $\lambda^f$, then

$$\tau^{a'}_{b'\,c'}\,X^{c'}_f = X^{a'}_d\, X^e_{b'}\, \tau^d_{e\,f}\,. \tag{1.77}$$

Thus,

$$\tau^{a'}_{b'\,h'} = \tau^{a'}_{b'\,c'}\,\delta^{c'}_{h'} \overset{(1.69)}{=} \tau^{a'}_{b'\,c'}X^{c'}_f\,X^f_{h'} \overset{(1.77)}{=} X^{a'}_d\, X^e_{b'}\, X^f_{h'}\,\tau^d_{e\,f}\,,$$

which transforms according to equation (1.75) and, so, $\boldsymbol{\tau}$ is a tensor. ■

Definition The **Kronecker tensor** is a symbol $\boldsymbol{\kappa} = \kappa^a_b = \delta^a_b$ whose components are Kronecker deltas.

Theorem The Kronecker tensor is a tensor.

Proof $X^{a'}_c\, X^d_{b'}\,\kappa^c_d = X^{a'}_c\, X^d_{b'}\,\delta^c_d = X^{a'}_c\, X^c_{b'} = \delta^{a'}_{b'} = \kappa^{a'}_{b'}$ ■

Note that

$$\kappa^{a'}_{b'} = \delta^{a'}_{b'} \overset{(*)}{=} \delta^{a}_{b} = \kappa^{a}_{b}\ .$$

This shows that $\boldsymbol{\kappa}$ is coordinate-independent. Because the components never change, it is customary to write $\delta^a_b$ for the components instead of $\kappa^a_b$.

The step $\delta^{a'}_{b'} \overset{(*)}{=} \delta^a_b$, above, is true because $a'$ and $b'$ range from 1 to $N$, and so do $a$ and $b$. Thus, the two Kronecker deltas have the same meaning:

$$\delta^{\text{top}}_{\text{bot}} = \begin{cases} 1 & \text{if top} = \text{bot} \\ 0 & \text{otherwise} \end{cases}.$$

Exercise 1.8.1 Show that there is no type (0,2) or (2,0) analog of a Kronecker tensor that is coordinate-independent.

Solution. Suppose $\boldsymbol{\tau} = \tau_{a\,b}$ is a type (0,2) tensor such that $\tau_{a\,b} = \delta_{a\,b}$. To disprove a conjecture normally requires a counter-example. However, we will instead show that $\tau_{a'\,b'} \neq \delta_{a'\,b'}$ for any coordinate transformation (other than the identity transformation).

The coordinate transformation for a (0,2) tensor is given by equation (1.8-3) :

$$\tau_{a'\,b'} = X^c_{a'}\, X^d_{b'}\, \tau_{c\,d} = X^c_{a'}\, X^d_{b'}\, \delta_{c\,d}.$$

To get the matrix version of this equation, let $\dot{\boldsymbol{T}}$ represent the primed matrix $(\tau_{a'\,b'})$ and $\boldsymbol{T}$ the unprimed matrix $(\tau_{c\,d})$. Since $\tau_{a\,b} = \delta_{a\,b}$, then $\boldsymbol{T} = \boldsymbol{I}$. Next, $X = (\frac{\partial x^{a'}}{\partial x^c})$ is the Jacobian transformation matrix, and so $X^{-1}$ is the matrix that represents both $X^c_{a'}$ and $X^d_{b'}$. Thus, we get

$$\dot{\boldsymbol{T}} = X^{-1} \boldsymbol{I}\, X^{-1} = \left(X^{-1}\right)^2 \neq \boldsymbol{I}.\ \text{That is, } \tau_{a'\,b'} \neq \delta_{a'\,b'} \qquad \blacksquare$$

By way of comparison, $\boldsymbol{\kappa}$ *is* coordinate-independent is because the coordinate transformation for a (1,1) tensor is given by equation (1.8-4),

$$\kappa^{a'}_{b'} = X^{a'}_{c}\, X^{d}_{b'}\, \kappa^{c}_{d},$$

and, thus the matrix representation is

$$\dot{\boldsymbol{K}} = X\, \boldsymbol{I}\, X^{-1} = X\, X^{-1} = \boldsymbol{I}.\ \text{That is, } \kappa^{a'}_{b'} = \delta^{a'}_{b'}.$$

Definition A tensor has **symmetry** with respect to a pair of indices if the tensor is unchanged when the indices are exchanged. For Type (0,2) this means $\tau_{a\,b} = \tau_{b\,a}$ for all $a$ and $b$. For Type (2,0) this means $\tau^{a\,b} = \tau^{b\,a}$ for all $a$ and $b$.

Example The Kronecker tensor $\boldsymbol{\kappa}$ is a type (1,1) example of a tensor with symmetry because $\boldsymbol{\kappa} = \delta^a_b = \delta^b_a$ for all $a$ and $b$.

Definition A tensor has **antisymmetry** (or **skew symmetry**) if the tensor changes sign when the indices are exchanged. For Type (0,2) this means $\tau_{a\,b} = -\tau_{b\,a}$ for all $a$ and $b$. For Type (2,0) this means $\tau^{a\,b} = -\tau^{b\,a}$ for all $a$ and $b$.

Theorem 1.8.2 (Exercise 1.8.2) Tensor symmetry and skew symmetry are coordinate-independent.

Proof. Given that $\tau^{a\,b} = \tau^{b\,a}$, we wish to show $\tau^{a'\,b'} = \tau^{b'\,a'}$.

$$\tau^{a'\,b'} \overset{(1.74)}{=} X^{a'}_c\, X^{b'}_d\, \tau^{c\,d} = X^{a'}_c\, X^{b'}_d\, \tau^{d\,c} = X^{b'}_d\, X^{a'}_c\, \tau^{d\,c} \overset{(1.74)}{=} \tau^{b'\,a'} \qquad \blacksquare$$

We now turn our attention to metric tensors. In this section, metric tensors cannot be defined, simply, as $g_{a\,b} \equiv \mathbf{e}_a \cdot \mathbf{e}_b$ and $g^{a\,b} \equiv \mathbf{e}^a \cdot \mathbf{e}^b$ because we have chosen not to introduce bases. A more round-about approach is necessitated.

Moreover, not every differentiable manifold has metric tensors. For the remainder of this book, we will restrict our attention to manifolds that do have metric tensors.

Definition A **covariant metric tensor** is a symmetric (0,2) tensor $\mathbf{g} = g_{a\,b}$ whose matrix $G = (g_{a\,b})$ is invertible and has inverse matrix $G^{-1} = (g^{a\,b})$ . In terms of components, these two conditions are

$$g_{a\,b} = g_{b\,a} \quad \text{and} \quad g^{a\,b} g_{b\,c} = \delta^a_c = \delta^c_a = g_{a\,b}\, g^{b\,c}. \tag{1.78}$$

An object having components $g^{a\,b}$ will be called a **contravariant metric tensor.** The label “tensor” for $g^{a\,b}$ is justified by the next theorem (which involves a convoluted proof because we have chosen not to introduce bases).

Convention Henceforth we will **refer to tensor components**, like $\tau^{a\,b}$**, as tensors**. We will **refer to vector components**, like $\lambda^a$, **as vectors**.

Theorem 1.8.3 $g^{ab}$ is a type (2,0) tensor.

Proof. I suggest the following simpler proof than the one in the book.
Since $g_{ab}$ is a tensor, it transforms as

$$g_{a'b'} = X^c_{a'} X^d_{b'} g_{cd} .$$

Since $(g_{a'b'})^{-1} = g^{a'b'}$, $(g_{cd})^{-1} = g^{cd}$, $\left(X^c_{a'}\right)^{-1} = X^{a'}_c$, and $\left(X^d_{b'}\right)^{-1} = X^{b'}_d$,
then $g^{a'b'}$ transforms as

$$g^{a'b'} = (g_{a'b'})^{-1} = (X^c_{a'} X^d_{b'} g_{cd})^{-1} = X^{a'}_c X^{b'}_d g^{cd} \qquad ■$$

Reminder Metric tensors, like all tensors, are defined at each point of a manifold. There is not generally a single metric tensor that represents the entire manifold.

Theorem 1.8.4 Metric tensors are not generally coordinate-independent.

Proof. Set $g_{ab} \equiv \tau_{ab} = \delta_{ab}$. By Exercise 1.8.1, $g_{ab}$ is not coordinate dependent for any non-trivial coordinate change. ■

We saw in Theorem 1.2.4 that the Euclidean metric tensors can be used to raise and lower indices in Euclidean vectors. This is also true for manifold metric tensors.

In the next example we show how **$g_{ab}$ can be used to lower an index of a tensor, and $g^{ab}$ can be used to raise an index,** by using contraction.

Definition Tensors that can be obtained from each other by raising or lowering indices are said to be **associated**. It is customary to use the same kernel letter for associated tensors, regarding them as different versions of the same tensor rather than as different tensors.

The following example, where superscript $c$ is lowered, illustrates this process.

Example Given a tensor $\tau^a$, define a new tensor $\tau_a$ as

$$\tau_a \equiv g_{ac}\tau^c = g_{a1}\tau^1 + ... + g_{an}\tau^n.$$

$\tau_a$ is a tensor because it is created by the contraction of two tensors. $\tau_a$ is associated with $\tau^a$ because it is formed by lowering the index ‘a’ .

Because $g_{a\,b}$ is symmetric, $\tau_a$ can also be written as $\tau_a = g_{c\,a}\tau^c$.

Take care, for example, that $g_{a\,1}\tau^1 \neq \tau_a$. One cannot lower superscript "1" because the common index must be a variable. The metric tensor $g_{a\,b}$ only lowers an index via *contraction*, and contraction requires a summation, as illustrated above. There is no summation in $g_{a\,1}\tau^1$.

Suppose we are given a contravariant tensor, $\tau^{a\,b}$. There are two ways to lower indices in $\tau^{a\,b}$; lower the first index or lower the second index:

1. $\boxed{\tau_a{}^b \equiv g_{a\,c}\,\tau^{c\,b}}$ (1.8-5)

2. $\boxed{\tau^a{}_b \equiv \tau^{a\,d}\,g_{d\,b}}$ (1.8-6)

In general, $\tau_a{}^b \neq \tau^a{}_b$ which is why we use spacing to distinguish them.

We can lower both indices.

$$\boxed{\tau_{a\,b} \equiv g_{a\,c}\,\tau^{c\,d}\,g_{d\,b}} \tag{1.8-7}$$

To see this,

$$g_{a\,c}\,\tau^{c\,d}\,g_{d\,b} \overset{(1.8\text{–}5)}{=} \tau_a{}^d\,g_{d\,b} = \tau_{a\,b} \quad \checkmark$$

and, also

$$g_{a\,c}\,\tau^{c\,d}\,g_{d\,b} \overset{(1.8\text{–}6)}{=} g_{a\,c}\tau^c{}_b = \tau_{a\,b} \quad \checkmark$$

That is, even though $\tau_a{}^b \neq \tau^a{}_b$, we can get from $\tau^{a\,b}$ to $\tau_{a\,b}$ via either $\tau_a{}^b$ or $\tau^a{}_b$.

There are also two ways to raise an index in $\tau_{a\,b}$.

1. $\boxed{\tau^a{}_b = g^{a\,c}\,\tau_{c\,b}}$ and $\boxed{\tau_a{}^b \equiv \tau_{a\,d}\,g^{d\,b}}$ (1.8-8)

and we can raise both indices:

$$\boxed{\tau^{a\,b} \equiv g^{a\,c}\,\tau_{c\,d}\,g^{d\,b}} \tag{1.8-9}$$

No matter which of the two routes we use to get from $\tau^{a\,b}$ to $\tau_{a\,b}$, and which of the two routes we use to get back, the resulting $\tau^{a\,b}$ is the same tensor we started with.

Example Suppose $(g^{ab}) = \begin{pmatrix} .5 & 0 \\ 0 & 2 \end{pmatrix}$ and $(\tau^{ab}) = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$. Find the three tensors associated with $\tau^{ab}$. Show that $\tau^a{}_b$ formed by lowering $\tau^{ab}$ is the same $\tau^a{}_b$ formed by raising $\tau_{ab}$, and similarly for $\tau_a{}^b$.

$$(g_{ab}) = \left(g^{ab}\right)^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix}.$$

$$(\tau_{ab}) = (g_{ac})\,(\tau^{cd})\,(g_{db}) = \begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 2 & -.25 \end{pmatrix} \checkmark$$

$$(\tau^a{}_b) = (\tau^{ad})\,(g_{db}) = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 4 & -.5 \end{pmatrix}$$

$$(\tau^a{}_b) = (g^{ac})\,(\tau_{cb}) = \begin{pmatrix} .5 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ 2 & -.25 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 4 & -.5 \end{pmatrix} \qquad \checkmark$$

$$(\tau_a{}^b) = (g_{ac})\,(\tau^{cb}) = \begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ 1 & -.5 \end{pmatrix}$$

$$(\tau_a{}^b) = (\tau_{ad})\,(g^{db}) = \begin{pmatrix} 4 & -2 \\ 2 & -.25 \end{pmatrix}\begin{pmatrix} .5 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ 1 & -.5 \end{pmatrix} \qquad \checkmark$$

Note that $\tau^a{}_b \neq \tau_a{}^b$. Also, raising and lowering of indices can be applied to any tensor.

Examples

$\tau^a{}_b{}^c = g_{bd}\,\tau^{adc}$ converts the rank (3,0) tensor $\tau^{abc}$ into the rank (2,1) tensor $\tau^a{}_b{}^c$

$\tau_a{}^b{}_c = g_{ad}\,\tau^{db}{}_c$ converts the rank (2,1) tensor $\tau^{ab}{}_c$ into the rank (1,2) tensor $\tau_a{}^b{}_c$

$\tau^a{}_{bc} = \tau^{ad}{}_c\,g_{db}$ converts the rank (2,1) tensor $\tau^{ab}{}_c$ into the rank (1,2) tensor $\tau^a{}_{bc}$

Theorem 1.8.5 $g^{ab}$, $g_{ab}$, and $\delta^a_b$ are associated.

Proof. $g^{ab} = g^{ac}\delta^b_c$ shows that raising index $c$ in $\delta^b_c$ gives $g^{ab}$. Thus, $g^{ab}$ and $\delta^a_b$ are associated. Similarly, $g_{ab}$ and $\delta^a_b$ are associated. $\therefore$ $g^{ab}$ and $g_{ab}$ are associated. ■

Note. The Kronecker delta is the one tensor for which we do not preserve the same kernel letter. As seen, raising $b$ in $\delta^a_b$ gives $g^{ab}$, not $\delta^{ab}$.

Example If $\tau^{ab} = \tau^{ba}$ then it is not necessary to apply spacing:

$\tau_a{}^b = g_{ac}\,\tau^{cb} = \tau^{bc}\,g_{ca} = \tau^b{}_a$ , and we can write $\tau^b_a = g_{ca}\,\tau^{cb}$ $\checkmark$

## 1.9 Metric Properties

We begin with a discussion of inner products.

Definition  Let $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ be two vectors in a manifold M. The **inner product** of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ is defined to be $g_{ab}\,\lambda^a\,\mu^b$.

Theorem 1.9.1  There are four ways of writing the inner product between vectors $\lambda^a$ and $\mu^a$:

$$\boxed{g_{ab}\,\lambda^a\,\mu^b = g^{ab}\,\lambda_a\,\mu_b = \lambda_a\,\mu^a = \lambda^a\,\mu_a} \tag{1.79}$$

Proof. $g_{ab}$ lowers $\lambda^a$: $g_{ab}\lambda^a = \lambda_b$. So $g_{ab}\,\lambda^a\,\mu^b = \lambda_b\,\mu^b$ ✓
$g_{ab}$ lowers $\mu^b$: $g_{ab}\mu^b = \mu_a$. So $g_{ab}\,\lambda^a\,\mu^b = \lambda^a\mu_a$ ✓
$g^{ab}$ raises $\mu_b$: $g^{ab}\,\mu_b = \mu^a$. So $\mathrm{g}^{ab}\,\lambda_a\,\mu_b = \lambda_a\,\mu^a$ ✓ ■

Definition  The **Euclidean dot product** is defined as $\boldsymbol{\lambda} \bullet \boldsymbol{\mu} \equiv \lambda^a \mathbf{e}_a \bullet \mu^b \mathbf{e}_b$.

Since $\mathbf{e}_a \bullet \mathbf{e}_b \overset{(1.22)}{=} g_{ab}$, then $\boldsymbol{\lambda} \bullet \boldsymbol{\mu} = g_{ab}\,\lambda^a\,\mu^b$. This shows that the Euclidean dot product for a manifold is an inner product.

We will see in Appendix A that the **spacetime dot product** must be defined differently.

Definitions
- An **inner product is positive definite** if for all vectors $\boldsymbol{\lambda}$
  $g_{ab}\,\lambda^a\lambda^b \geq 0$, and $g_{ab}\,\lambda^a\lambda^b = 0$ iff $\boldsymbol{\lambda} = 0$.
- An inner product that is not positive definite is an **indefinite inner product**.
- A manifold that has a positive definite inner product at every point is called a **Riemannian manifold.** We say that a Riemannian manifold has **a positive definite metric tensor field**.
- A manifold that possesses an indefinite metric tensor field is called a **pseudo-Riemannian manifold** or a **semi-Riemannian manifold**.

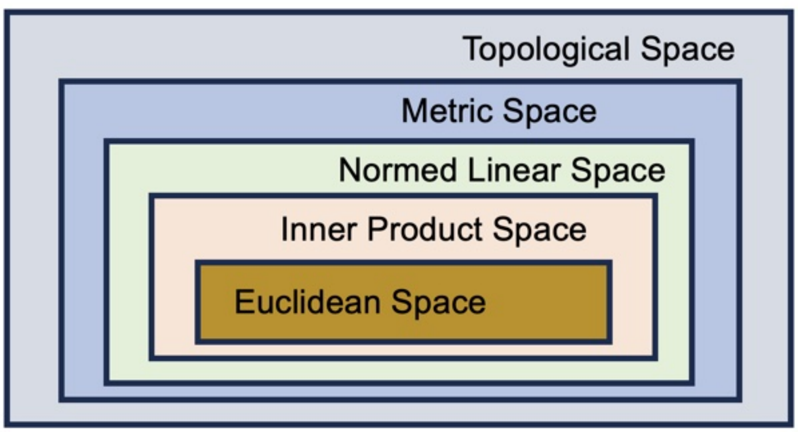
Using this terminology, the Euclidean dot product is a positive definite inner product. We use it to define Euclidean $N$-space. First, recall from Section 1.8 that $\mathbb{R}^N$ is the set of points $\left(x^1, x^2, \ldots, x^N\right)$ having real coordinates. It becomes a vector space under the operations of coordinate-wise addition and scalar multiplication by real numbers.

Definitions

- **Euclidean *N*-space $\mathbb{E}^N$** consists of the vector space $\mathbb{R}^N$ endowed with the dot product for its inner product.
- A **pseudo Euclidean *N*-space** consists of the vector space $\mathbb{R}^N$ endowed with an indefinite inner product.

Definition A **pseudo metric** is a "metric" that allows the distance between distinct points to be zero. A **pseudo norm** is a "norm" that allows a non-zero vector to have zero length.

The next theorem applies to both Euclidean and pseudo Euclidean *N*-spaces. It gives the relationship of the topological spaces used in these definitions, shown in the Venn diagram below. It shows the path from $\mathbb{E}^N$ to $\mathbb{R}^N$. The symbol $\mathbb{R}^N$ has been used for both the topological space and the metric (a.k.a., linear) space.




Theorem For all *N*,

Euclidean *N*-space ⊂ Inner Product Space ⊂ Normed Linear Space ⊂ Metric space ⊂Topological Space

Proof.

1. Euclidean space is endowed with the dot product, which is an inner product
2. An inner product generates a norm via the formula

$$\boxed{\|\boldsymbol{\lambda}\| = \sqrt{\left| g_{ab}\, \lambda^a \lambda^b \right|}} \tag{1.80}$$

3. A norm generates a metric via the formula

$$d(\boldsymbol{\lambda}, \boldsymbol{\mu}) \equiv \|\boldsymbol{\lambda} - \boldsymbol{\mu}\|.$$

4. A metric generates open balls of radius $r$ about every point P:
   $B_r(\mathrm{P}) = \{\mathrm{Q}: \mathrm{d}(\mathrm{P},\mathrm{Q}) < r\}$.
   The open balls generate the rest of the topology (i.e., the remaining open sets) using the rules:
   Any union of open sets is open
   Any finite intersection of open sets is open. ■

With this definition of norm, we have that inner products are positive definite when non-zero vectors have non-zero length. Thus, flat Euclidean spaces are positive definite because $\|\boldsymbol{\lambda}\| = \sqrt{\boldsymbol{\lambda}\cdot\boldsymbol{\lambda}} = 0$ iff $\boldsymbol{\lambda} = 0$. But, pseudo Euclidean spaces, like the flat space-time of Special Relativity, have at least one non-zero vector $\boldsymbol{\lambda}$ with length $\|\boldsymbol{\lambda}\| = 0$, and it is possible to have a non-trivial arc segment with zero arc length,
$\int_\gamma \sqrt{g_{a b}\, \dot{x}^a \dot{x}^b}\, \mathrm{d}t = 0$.

To accommodate pseudo-Riemannian manifolds like spacetime, we make the following definitions.

Definition A **non-zero vector $\boldsymbol{\lambda}$ is** $\begin{pmatrix} \textbf{timelike} \\ \textbf{null}, \text{ or } \textbf{lightlike} \\ \textbf{spacelike} \end{pmatrix}$ if $g_{\mu\nu}\, \lambda^\mu \lambda^\nu \begin{pmatrix} > 0 \\ = 0 \\ < 0 \end{pmatrix}$.

A **curve is** $\begin{pmatrix} \textbf{timelike} \\ \textbf{null}, \text{ or } \textbf{lightlike} \\ \textbf{spacelike} \end{pmatrix}$ if the tangent vectors are $\begin{pmatrix} \text{timelike} \\ \text{null, or lightlike} \\ \text{spacelike} \end{pmatrix}$.

If, say, ‘timelike’ applies to only part of a curve, we describe just that part as timelike.

We extend the definitions of Riemannian metric properties by taking absolute values of inner products. We have already defined (in equation 1.80) the **length of a vector $\boldsymbol{\lambda}$** as

$$L = \sqrt{\left|g_{a b}\, \lambda^a \lambda^b\right|} = \sqrt{\left|g^{a b}\, \lambda_a \lambda_b\right|} = \sqrt{|\lambda_a \lambda^a|}$$

A **unit vector** is a vector whose length is one. A **null vector** is a non-zero vector whose length is zero.

The **angle between two non-null vectors** is given by

$$\boxed{\cos\theta = \frac{g_{ab}\,\lambda^a\,\mu^b}{\sqrt{g_{cd}\,\lambda^c\,\lambda^d}\ \sqrt{g_{ef}\,\mu^e\,\mu^f}}}\,. \tag{1.81}$$

**Vectors $\lambda$ and $\mu$ are orthogonal** if $g_{ab}\,\lambda^a\,\mu^b = 0$. Thus, null vectors are orthogonal to themselves.

If $\gamma = \{x^a(t) : a \leq t \leq b\}$ is a curve in a manifold M, then the **length of $\gamma$** is

$$\boxed{L = \int_a^b ds = \int_a^b \sqrt{\left|\,g_{ab}\,\dot{x}^a\,\dot{x}^b\,\right|}\ dt} \tag{1.82}$$

and the infinitesimal length of its **line element** is

$$\boxed{ds^2 = \left|g_{ab}\,dx^a\,dx^b\right|}\,. \tag{1.83}$$

A **null curve** is one with $g_{ab}\,\dot{x}^a\,\dot{x}^b = 0$ at every point $x^a \in \gamma$.

Exercise 1.9.2 The length of $\gamma$ is well-defined. That is, the definition of curve length is coordinate-independent.

Exercise 1.9.2 Equation (1.81) results in $|\cos\theta| \leq 1$.

Two points could be connected by curves with different length. Thus, we only define curve length, not the distance between points.

Definition **A metric tensor field** $g_{\mu\nu}$ ($\mu$, $\nu$ = 0, 1, 2, 3) on a manifold M **has indefiniteness (+---)** means that at every point P ∈ M, a coordinate system that leads to G = ($g_{\mu\nu}$) being a diagonal matrix has a positive first diagonal element while the other three are negative. Physicist generally use indefiniteness (+---) for spacetime. Mathematicians usually use (-+++).

## 1.10 What and where are the bases?

In Euclidean 3-space $\left(\mathbb{E}^3\right)$ we started with a Cartesian coordinate system $(x^i)$, postulated an alternate coordinate system $(u^i)$, and then defined natural basis vectors (tangents) and dual basis vectors (gradients):

$$\mathbf{e}_i \overset{(1.2\text{–}1)}{=} \frac{\partial \boldsymbol{r}}{\partial u^i} = \frac{\partial x}{\partial u^i}\boldsymbol{i} + \frac{\partial y}{\partial u^i}\boldsymbol{j} + \frac{\partial z}{\partial u^i}$$

$$\mathbf{e}^i \overset{(1.2\text{–}2)}{=} \nabla u^i = \frac{\partial u^i}{\partial x}\boldsymbol{i} + \frac{\partial u^i}{\partial y}\boldsymbol{j} + \frac{\partial u^i}{\partial z}\boldsymbol{k} \quad .$$

A manifold M begins with a locally Euclidean coordinate system $(x^a)$ at each point P, and we can postulate an alternate coordinate system $(u^a)$. But, we cannot define

$$\mathbf{e}_a = \frac{\partial x^1}{\partial u^a}\boldsymbol{i} + \frac{\partial x^2}{\partial u^a}\boldsymbol{j} + \frac{\partial x^3}{\partial u^a}\boldsymbol{k} + \frac{\partial x^4}{\partial u^a}\boldsymbol{l} + \ldots + \frac{\partial x^N}{\partial u^a}\boldsymbol{m}$$

$$\mathbf{e}^a = \frac{\partial u^a}{\partial x^1}\boldsymbol{i} + \frac{\partial u^a}{\partial x^2}\boldsymbol{j} + \frac{\partial u^a}{\partial x^3}\boldsymbol{k} + \frac{\partial u^a}{\partial x^4}\boldsymbol{l} + \ldots + \frac{\partial u^a}{\partial x^N}\boldsymbol{m} \, .$$

The ***i*-*j*-*k*** paradigm works only in 3 dimensions, primarily because cross products have no meaning outside of 3-space. Rather, what works is to define the bases in terms of the coordinate curves and the level surfaces, using Euclidean coordinates.

If we allow only the *b*th coordinate to vary while keeping all others fixed, we obtain a coordinate curve in M that can be parameterized as

$$x^a = x^a_{\mathrm{P}} + \delta^a_b\, t,$$

where P = $(x^1_{\mathrm{P}} \ldots x^b_{\mathrm{P}} \ldots x^N_{\mathrm{P}})$. That is, P($t$) = $(x^1_{\mathrm{P}} \ldots x^b_{\mathrm{P}} + t \ldots x^N_{\mathrm{P}})$. Clearly,

$$\frac{\mathrm{d}x^a}{\mathrm{d}t} = \delta^a_b.$$

The **tangent vector** vector $\left(\frac{\mathrm{d}x^a}{\mathrm{d}t}\right)$ is what we desire for the contravariant basis vector $\mathbf{e}_a$ , but its indices are superscripts, not the expected subscripts. This is not a fundamental problem, just a matter of expressing the equation using tensor notation. Foster and Nightengale define $\mathbf{e}_a$ to be $\frac{\mathrm{d}x^a}{\mathrm{d}t}$ using words in order to avoid the equation. Nonetheless, I give the equation definition, below, in order not to ignore this problem. In Appendix D.3, I have generated a more modern development for basis $\mathbf{e}_a$ that avoids this problem. There is no Appendix D in the book.

Definition The **contravariant basis** is composed of the tangent vectors

$$\mathbf{e}_a \equiv (\delta_a^b) = \left(\frac{dx^a}{dt}\right) = (0, \ldots, 1_a, \ldots, 0). \tag{1.10-1}$$

For the covariant vectors, recall definition (1.7-4) that a function $\varphi$ that picks out the $a$th coordinate of each point of M is an example of a scalar field.

Since P = ( $x_P^1 \ldots x_P^a \ldots x_P^N$) = ($x_P^a$), then

$$\varphi(x_P^a) \equiv x_P^a.$$

By definition (1.7-5), the **gradient of** $\varphi$ is the covariant vector $\nabla\varphi$ having components $\partial_b\varphi$. This gradient evaluated at P, $\nabla\varphi(x^a)$, has components $\partial_b\varphi(x^a)$; i.e.,

$$\nabla\varphi(x^a) = \left(\partial_b\varphi(x^a)\right) = \left(\frac{\partial x^a}{\partial x^b}\right) = \left(\delta_b^a\right).$$

We make this the definition of the covariant basis.

Definition The **covariant basis** is composed of the gradient vectors

$$\mathbf{e}^a \equiv \left(\delta_b^a\right) = \left(\frac{\partial x^a}{\partial x^b}\right) = (0, \ldots, 1^a, \ldots, 0) \tag{1.10-2}$$

Definition The vector space at a point P∈M generated by the tangent basis is called the **tangent space**, $\mathbf{T}_P = \{\lambda^a \mathbf{e}_a\}$, and the vector space generated by the gradient basis (1-forms) is called the **cotangent space**, designated $\mathbf{T}_P{}^* = \{\mu_b \mathbf{e}^b\}$.

Just as in a Euclidean manifold where the vectors and dual vectors do not lie in the manifold, we consider $\mathbf{T}_P$ and $\mathbf{T}_P{}^*$ to be attached to M at the point P but not part of M. Note that M is not necessarily embedded in any larger space where tangent and cotangent planes could lie.

It might seem at first that this approach yields $\mathbf{e}_a = \mathbf{e}^a$ because they have the same components, all zeros except that the $a$th component is 1. In point of fact, we regard $\mathbf{e}_a$ to be different from $\mathbf{e}^a$ because they are different types of objects, contravariant and covariant. This is exactly analogous to how we regard row and column vectors having the same components to be different – they are different kinds of objects and, as such, are manipulated somewhat differently.

Contrast this with Chapter 1 where we treated contravariant and covariant vectors as though they were the same kind of object: both were Euclidean vectors $ai + bj + ck$. They could sometimes be equal. In fact, that was not quite true but we we able to get away with it because we were using the Euclidean metric tensor $g_{ij} = \delta_{ij}$ which gave us $\mathbf{e}_i = g_{ij}\, \mathbf{e}^j = \delta_{ij}\, \mathbf{e}^j = \mathbf{e}^i$. However, this only means that their components are equal, not that the basis vectors themselves are equal. In curved space, $g_{ij} \neq \delta_{ij}$, and we can't get away with treating them as the same kind of object.

Lastly, we discuss the Euclidean dot product and its tie to the metric tensors. Recall that the metric tensor $g_{ab}$ is a symmetric tensor that is invertible. Extending the Euclidean 3-space definition of dot product, we see that if we set

$$g_{ab} = \boldsymbol{e}_a \bullet \boldsymbol{e}_b = \begin{pmatrix} 0 \\ \vdots \\ 1_a \\ \vdots \\ 0 \end{pmatrix} \bullet \begin{pmatrix} 0 \\ \vdots \\ 1_b \\ \vdots \\ 0 \end{pmatrix} = 0 + 0 + \ldots + 0 + \delta_{ab} + 0 + \ldots + 0 = \delta_{ab}$$

and

$$g^{ab} = \boldsymbol{e}^a \bullet \boldsymbol{e}^b = (0 \;\ldots\; 1^a \;\ldots\; 0) \bullet (0 \;\ldots\; 1^b \;\ldots\; 0) = \delta^{ab},$$

then $g_{ab}$ is a metric tensor because it is symmetric ($g_{ba} = g_{ab}$) and invertible ($g_{ab}\, g^{bc} = \mathrm{I}$, the identity matrix).

We have identified the bases for *vectors*. Formal development of the bases for *tensors*, is deferred until Appendix C.3 because it depends on an understanding of dual spaces. The key concepts are the tensor product of basis vectors (C3-5) and the general tensor definition (C3-9).

# Chapter 2 (part 1) Geodesics and Differentiation

## 2.0 Introduction

In flat Euclidean 3-space, $\mathbb{E}^3$, a geodesic is a straight line. It can be characterized by the fact that its tangent vectors are parallel. We will call this characterization “parallel transport” and think of it as transporting one tangent vector to each of the other points on the line.

To underscore the similarity between geodesics and parallel transport, the defining equation (2.3) in $\mathbb{E}^3$ for a geodesic is identical to equation (2.22) in $\mathbb{E}^3$ for parallel transport. Using index notation, and using dot notation for differentiation, both equations read

$$\dot{\lambda}^i + \Gamma^i_{j\,k}\, \lambda^j\, \dot{u}^k = 0\ .$$

The difference is that the vector $\boldsymbol{\lambda}$ in equation (2.3) is a tangent vector to a curve that is parallel to all the other tangent vectors, while the vector $\boldsymbol{\lambda}$ in the equation (2.22) is an arbitrary vector that is being transported.

In Section 2.1, we develop the concept of geodesics, first in flat Euclidean space, then in a curved manifold. Then, in Section 2.2, we do the same for the concept of parallel transport. For both geodesics and parallel transport, we use the derived Euclidean equations as the model for the manifold definitions. Consider for example the Euclidean and manifold equations, respectively, for parallel transport:

$$\dot{\lambda}^i + \Gamma^i_{j\,k}\, \lambda^j\, \dot{u}^k = 0 \qquad \text{and} \qquad \dot{\lambda}^a + \Gamma^a_{b\,c}\, \lambda^b\, \dot{x}^c = 0$$

This might suggest that parallel transport vectors in a manifold are just like parallel vectors in Euclidean space, and this is not true. Consider a sphere. It is a curved manifold embedded in flat Euclidean 3-space. A geodesic on a sphere is a great circle, which is not a geodesic (i.e., straight line) in $\mathbb{E}^3$. Similarly, vectors that are defined to be parallel on the sphere will not be parallel in $\mathbb{E}^3$. For example, the parallel tangent vectors on a great circle point in many different directions in $\mathbb{E}^3$.

While this may seem obvious now, it can get confusing when we delve into the details, and we need to be prepared to take a step back and reflect on the difference. We should remember that geodesics and parallel transport vectors defined on a manifold are not geodesics and parallel vectors in $\mathbb{E}^3$.

## 2.1 Geodesics

In Euclidean space, a geodesic is often characterized as the shortest path between two points. However, in spacetime, some curves have zero length because the tensor field is indefinite: that is, $g_{a\,b}\,\lambda^a\lambda^b = 0$ even though $\lambda^a \neq 0$. So, the property we use to define geodesic is that a curve is straight; i.e., all tangent lines to the curve point in the same direction. We begin by characterizing this in Euclidean space and then generalize to non-Euclidean space.

### Euclidean 3-Space

Notation  Recall from Section 1.2 for Euclidean space:

Index $i = 1, 2, 3$

$(x,y,z) = (x^i)$ is the Cartesian coordinate system

$(u,v,w) = (u^j)$ is an alternate coordinate system

$\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is the Cartesian basis

$\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$ is the position vector (1.3)

$\{\mathbf{e}^u, \mathbf{e}^v, \mathbf{e}^w\} = \{\mathbf{e}^i\}$ , $\{\mathbf{e}_u, \mathbf{e}_v, \mathbf{e}_w\} = \{\mathbf{e}_i\}$, and

$$\mathbf{e}_i \equiv \frac{\partial \boldsymbol{r}}{\partial u^i} = \frac{\partial x}{\partial u^i}\boldsymbol{i} + \frac{\partial y}{\partial u^i}\boldsymbol{j} + \frac{\partial z}{\partial u^i}\boldsymbol{k} \quad \text{is the natural basis} \tag{1.2-1}$$

$$\mathbf{e}^i \equiv \nabla u^i = \frac{\partial u^i}{\partial x}\boldsymbol{i} + \frac{\partial u^i}{\partial y}\boldsymbol{j} + \frac{\partial u^i}{\partial z}\boldsymbol{k} \quad \text{is the dual basis} \tag{1.2-2}$$

Definition  We say that **Euclidean space** (or a subspace of Euclidean space) **is parameterized with parameter *t*** if points can be represented parametrically:

$x = x(t)$

$y = y(t)$

$z = z(t)$ .

Suppose a curve $\gamma$ is parameterized by arc length, $s$. The tangent vectors to $\gamma$ can then be expressed by $\boldsymbol{\lambda} = \boldsymbol{\lambda}(s) \equiv \dfrac{d\boldsymbol{r}(s)}{ds}$. These tangent vectors $\boldsymbol{\lambda}(s)$ are unit vectors:

$$\boldsymbol{\lambda} = \frac{d\boldsymbol{r}}{ds} = \frac{dx}{ds}\boldsymbol{i} + \frac{dy}{ds}\boldsymbol{j} + \frac{dz}{ds}\boldsymbol{k} \Rightarrow |\,\boldsymbol{\lambda}\,|^2 = \left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = \left(\frac{ds}{ds}\right)^2 = 1 \quad \checkmark$$

The tangent vectors can also be expressed as $\boldsymbol{\lambda} \overset{1.16}{=} \lambda^i\, \boldsymbol{e}_i$ where

$$\lambda^i = \frac{\mathrm{d}u^i}{\mathrm{d}s}: \tag{2.1-1}$$

$$\lambda^i\, \mathbf{e}_i \overset{1.16}{=} \boldsymbol{\lambda} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} = \frac{\partial \boldsymbol{r}}{\partial u^i} \frac{\mathrm{d}u^i}{\mathrm{d}s} \overset{1.2-1}{=} \frac{\mathrm{d}u^i}{\mathrm{d}s} \mathbf{e}_i \quad ✓$$

A vector has magnitude and direction. Differentiation of a vector operates on both the magnitude and the direction. If the magnitude does not vary, then the derivative depends only on the direction. If we use arc length as the parameter, the vectors $\boldsymbol{\lambda}(s)$ are unit vectors. So, their derivatives depend only on direction. For a curve to be a straight line, all the tangents must point in the same direction:

$$\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}s} = 0. \tag{2.1}$$

The next theorem characterizes parameters of straight lines that satisfy equation (2.1).

Theorem 2.1.1 Let a straight line be parameterized by $t$ and let $\boldsymbol{\mu} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$ be the tangent vector. Then $\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = 0$ iff there are integers $A$ and $B$, $A \neq 0$, such that $t = A\, s + B$. Moreover, if $t = A\, s + B$ then all vectors $\boldsymbol{\mu}(t)$ have the same length.

Proof. Let $s = f(t)$. Then $\boldsymbol{\mu} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} \frac{\mathrm{d}s}{\mathrm{d}t} = \boldsymbol{\lambda}\, f'(t)$. So,

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}t} f'(t) + \boldsymbol{\lambda}\, f''(t) = \frac{d\boldsymbol{\lambda}}{\mathrm{d}s} \frac{\mathrm{d}s}{\mathrm{d}t} f'(t) + \boldsymbol{\lambda}\, f''(t) \overset{2.1}{=} \boldsymbol{\lambda}\, f''(t),$$

where, as discussed above, the tangent vector $\boldsymbol{\lambda} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s}$ has unit length. Thus,

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = 0 \iff f''(t) = 0 \iff f'(t) = C \iff s = f(t) = C\, t + D \iff t = A\, s + B$$

where $A = \frac{1}{C}$ and $B = -\frac{D}{C}$. (Note that $C \neq 0$ because otherwise $C$ would be a constant, not a parameter).

Finally, if $t = A\, s + B$ then $|\boldsymbol{\mu}(t)| = |\boldsymbol{\lambda}(t)\, f'(t)| = |\boldsymbol{\lambda}(t)| \left| \frac{\mathrm{d}s}{\mathrm{d}t} \right| = 1 \cdot C = C$. ∎

Example Let the line $x = y = z$ be parameterized by $t = \sqrt{s}$ . Clearly, $t \neq A\, s + B$. So, by Theorem 2.1.1, if $\boldsymbol{\mu} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$ then $\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} \neq 0$. To confirm this, note that arc length on this diagonal line is $s^2 = x^2 + y^2 + z^2 = 3\, x^2$, or

$$x = \frac{s}{\sqrt{3}} \; .$$

Moreover,

$$s = t^2,$$

$$\boldsymbol{r} = x\, \boldsymbol{i} + y\, \boldsymbol{j} + z\, \boldsymbol{k} \; = x\, (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}) = \frac{s}{\sqrt{3}}\, (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}) = \frac{t^2}{\sqrt{3}}\, (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}).$$

Set $\boldsymbol{\lambda} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} = \frac{1}{\sqrt{3}}\, (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k})$ and $\boldsymbol{\mu} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{2\, t}{\sqrt{3}}\, (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k})$. Then

$$\frac{d\boldsymbol{\lambda}}{\mathrm{d}s} = 0$$

but

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \frac{2}{\sqrt{3}}\, (\boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}) \neq 0.$$ ■

Equation (2.1) is the geodesic equation. We re-express this in terms of the natural basis, equation (2.2); then in terms of tangent vector components and coefficients $\Gamma^k_{ij}$, equation (2.3); and, finally, purely in terms of the alternate coordinates, equation (2.4).

We begin with

$$0 \overset{(2.1)}{=} \frac{d\boldsymbol{\lambda}}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}(\lambda^i\, \mathbf{e}_i) = \frac{\mathrm{d}\lambda^i}{\mathrm{d}s}\mathbf{e}_i + \lambda^i\, \frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}s}, \tag{2.2}$$

which underscores that equation (2.1) consists of three simultaneous equations, one for each basis vector $\mathbf{e}_i$ .

Next, $\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}s} \overset{\text{Chain Rule}}{=} \frac{\partial \mathbf{e}_i}{\partial u^j}\frac{\mathrm{d}u^j}{\mathrm{d}s} = \frac{\mathrm{d}u^j}{\mathrm{d}s}\, \partial_j \mathbf{e}_i$ , where $\partial_j \equiv \frac{\partial}{\partial u^j}$. Plugging $\frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}s}$ into (2.2) gives

$$\frac{\mathrm{d}\lambda^i}{\mathrm{d}s}\mathbf{e}_i + \lambda^i\, \frac{\mathrm{d}u^j}{\mathrm{d}s}\, \partial_j \mathbf{e}_i = 0. \tag{2.1-2}$$

Since $\partial_j \mathbf{e}_i$ is a vector, we can express it in terms of the natural basis:

There are 27scalars $\Gamma^k_{ij}$ such that

$$\boxed{\partial_j \mathbf{e}_i = \Gamma^k_{ij}\, \mathbf{e}_k}\,. \tag{2.1-3}$$

Replacing $\partial_j \mathbf{e}_i$ in equation (2.1-2) yields

$$\frac{d\lambda^i}{ds}\mathbf{e}_i + \lambda^i \frac{du^j}{ds}\,\Gamma^k_{ij}\,\mathbf{e}_k = 0.$$

Relabeling indices i → j, j → k, k → i in the second term in order to match $\mathbf{e}_i$'s gives

$$\frac{d\lambda^i}{ds}\mathbf{e}_i + \lambda^j \frac{du^k}{ds}\,\Gamma^i_{jk}\,\mathbf{e}_i = 0.$$

$$\Rightarrow \quad \frac{d\lambda^i}{ds} + \lambda^j \frac{du^k}{ds}\,\Gamma^i_{jk} = 0 \quad \text{for } i = 1 - 3. \tag{2.3}$$

Recalling that $\lambda^i \overset{2.1-1}{=} \dfrac{du^i}{ds}$, equation (2.3) becomes

$$\boxed{\frac{d^2 u^i}{ds^2} + \Gamma^i_{jk}\frac{du^j}{ds}\frac{du^k}{ds} = 0} \quad \text{for } i = 1 - 3. \tag{2.4}$$

We next develop a formula to find the coefficients $\Gamma^i_{jk}$, recalling our convention in Section 1.7 that $x(u^i)$, $y(u^i)$, and $z(u^i)$ are analytic (so that the order of differentiation doesn't matter):

$$\partial_j\,\mathbf{e}_i \overset{(1.2-1)}{=} \frac{\partial^2 \boldsymbol{r}}{\partial u^j\,\partial u^i} = \frac{\partial^2 \boldsymbol{r}}{\partial u^i\,\partial u^j} = \partial_i\,\mathbf{e}_j \quad ✓$$

So,

$$\boxed{\Gamma^k_{ij} = \Gamma^k_{ji}}: \tag{2.5}$$

$$\Gamma^k_{ij}\,\mathbf{e}_k \overset{(2.1-3)}{=} \partial_j\,\mathbf{e}_i = \partial_i\,\mathbf{e}_j = \Gamma^k_{ji}\,\mathbf{e}_k \quad ✓$$

Recall that tensors are defined at a point P = $(u^i)$. So, though not obvious from the notation, the Euclidean metric tensor $g_{ij}$ is a function of the coordinates $(u^i)$. Thus, it makes sense to discuss operations like $\partial_k\, g_{ij}$.

Claim: $\partial_k\, g_{ij} = \Gamma^m_{ik}\, g_{mj} + \Gamma^m_{jk}\, g_{im}$ : (2.6)

$$\partial_k\, g_{ij} \overset{(1.22)}{=} (\partial_k \mathbf{e}_i)\cdot \mathbf{e}_j + \mathbf{e}_i \cdot (\partial_k \mathbf{e}_j) \overset{(2.1-3)}{=} \Gamma^m_{ik}\mathbf{e}_m \cdot \mathbf{e}_j + \mathbf{e}_i \cdot \Gamma^m_{jk}\mathbf{e}_m = \Gamma^m_{ik}\, g_{mj} + \Gamma^m_{jk}\, g_{im} \quad ✓$$

Relabeling (2.6) with $i \to j$, $j \to k$, $k \to i$ yields

$$\partial_i \, g_{jk} = \Gamma^m_{ji} \, g_{mk} + \Gamma^m_{ki} \, g_{jm} \tag{2.7}$$

Relabeling (2.7) with $i \to j$, $j \to k$, $k \to i$ yields

$$\partial_j \, g_{ki} = \Gamma^m_{kj} \, g_{mi} + \Gamma^m_{ij} \, g_{km} \tag{2.8}$$

Remembering that $g_{ij} = g_{ji}$ and $\Gamma^k_{ij} = \Gamma^k_{ji}$ , (2.6) + (2.7) – (2.8) gives

$$2 \, \Gamma^m_{ik} \, g_{mj} = \partial_k \, g_{ij} + \partial_i \, g_{jk} - \partial_j \, g_{ki} \, .$$

Contracting with $\frac{1}{2} \, g^{\ell j}$ yields

$$\boxed{\Gamma^\ell_{ik} = \frac{1}{2} \, g^{\ell j} \left(\partial_k \, g_{ij} + \partial_i \, g_{jk} - \partial_j \, g_{ki}\right)} : \tag{2.9}$$

$$\Gamma^\ell_{ik} = \Gamma^m_{ik} \, \delta^\ell_m = \Gamma^m_{ik} \, g^{\ell j} g_{mj} = \frac{1}{2} \, g^{\ell j} \left(\partial_k \, g_{ij} + \partial_i \, g_{jk} - \partial_j \, g_{ki}\right) \quad ✓$$

Note. Like $g_{ij}$, $\Gamma^i_{jk}$ is a function of $(u^\ell)$. So it will make sense (in later sections) to discuss operations like $\partial_\ell \, \Gamma^i_{jk}$.

Definition  Equation (2.4) with $\Gamma^i_{jk}$ given by equation (2.9) is known as the **geodesic equation for Euclidean space**. It represents the condition that the unit tangent vectors to a curve all point in the same direction.

As a result of Exercise 2.1.1, equation (2.4) is unchanged when replacing $s$ by

$$t = As + B, \;\; A \neq 0. \tag{2.10}$$

That is,

$$\boxed{\frac{d^2 u^i}{dt^2} + \Gamma^i_{jk} \frac{du^j}{dt} \frac{du^k}{dt} = 0} \quad \text{for all } i. \tag{2.11}$$

Definition  Euclidean parameters of the form $t = As + B$ are known as **affine parameters**. An **affine transformation** is a linear transformation ($t = As$) with a shift of origin (+ $B$). For an affine parameter, $\frac{ds}{dt}$ is constant so that movement along the geodesic is at a constant rate.

Theorem 2.1.2 Let a curve $\gamma$ have parameter $t$, and let $\boldsymbol{\mu} = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$ be the tangent vectors.

Parameter $t$ is affine $\Leftrightarrow$ Vectors $\boldsymbol{\mu}$ have constant length

$$\Leftrightarrow \frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = 0 \text{ is the equation of parallel vectors in } \mathbb{E}^3$$

Proof. $|\,\boldsymbol{\mu}\,| = \left| \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} \right| = \left| \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s} \right| \left| \frac{\mathrm{d}s}{\mathrm{d}t} \right| = (1) \left| \frac{\mathrm{d}s}{\mathrm{d}t} \right| = \left| \frac{\mathrm{d}s}{\mathrm{d}t} \right|$ is constant

iff $\left| \frac{\mathrm{d}s}{\mathrm{d}t} \right|$ is constant

iff $t$ is an affine transformation $As$ + B.

Since a vector has magnitude and direction, $\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = 0$ is the equation of vectors having the same direction iff $\boldsymbol{\mu}$ has constant length ■

The geodesic equation (2.11) along with (2.9) represents the condition that tangent vectors to an affinely parameterized curve all point in the same direction.

Solving equation (2.11) requires six conditions: three for $u^i$ and three for $\Gamma^i_{jk}$ (in equation 2.9). For example, the conditions could be (a) a starting point and a direction or (b) starting and ending points. That is, we can use these equations to generate a straight line from a point P in a specific direction, or we can generate a straight line from P to Q.

## *N*-Dimensional Riemannian and Pseudo Riemannian Spaces

The important property of *affine* parameters $t$ is not that they are linear combinations of of $s$ but that the tangent vectors (with respect to $t$) are parallel. In fact, if a curve $\gamma$ is null, then $s$ is constant and thus cannot even be a parameter. So, to generalize the Euclidean definition, we use equations (2.11 & 2.9) to define an affinely parameterized geodesic (equation 2.12), and then we derive (in Th 2.22) that the geodesic tangent vectors have constant length and, hence, are parallel.

In the development that follows, coordinate system $\{u^i\} = \{\mathrm{u,v,w}\}$ for Euclidean space is replaced by coordinate system $\{x^a\}$ at point P, curve parameter $t$ is replaced by $u$, and $\partial_i = \frac{\partial}{\partial u^i}$ is replaced by $\partial_a = \frac{\partial}{\partial x^a}$.

Recall that tensors are defined at points P = ($x^a$). So, though not transparent from the notation, the tensor $g_{ab}$ is a function of $x^a$, and so it makes sense to write expressions like $\partial_c g_{ab}$ as in (2.13), below.

Definition An **affinely parameterized geodesic** in an *N*-dimensional Riemannian or pseudo-Riemannian manifold is a curve $x^a(u)$ that satisfies

$$\boxed{\frac{d^2 x^a}{du^2} + \Gamma^a_{b\,c} \frac{dx^b}{du} \frac{dx^c}{du} = 0} \tag{2.12}$$

where the $N^3$ quantities $\Gamma^a_{b\,c}(u)$ are given in terms of $g^{ab}(u)$ and $g_{ab}(u)$ by

$$\boxed{\Gamma^a_{b\,c} \equiv \frac{1}{2} g^{ad} (\partial_b g_{dc} + \partial_c g_{bd} - \partial_d g_{bc})}\,. \tag{2.13}$$

The scalar functions $\boldsymbol{\Gamma^a_{b\,c}}$ are called the **connection coefficients.**

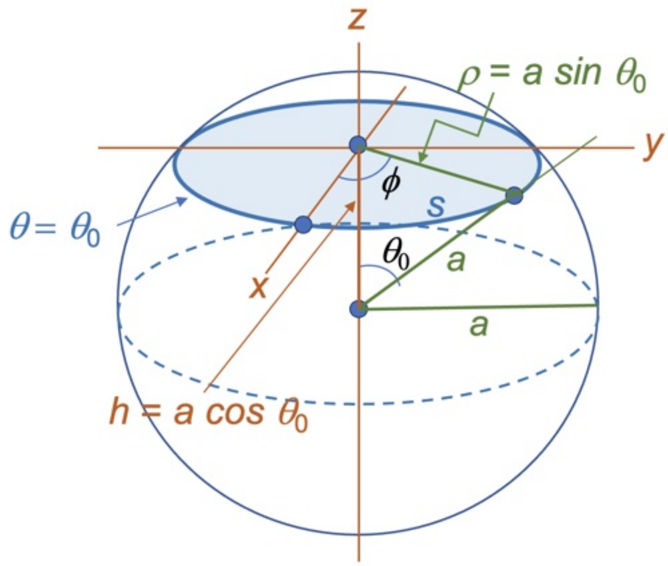
As pointed out in the Introduction, Section 1.0, whenever we extrapolate a Euclidean result into a spacetime definition we are required to show that our extrapolation is coordinate-independent. This will be deferred until Corollary 2.2.7 in Section 2.2.

Commutativity of subscripts *b* and *c* follows immediately from definition (2.13) since $g_{\text{ab}}$ is commutative.

Theorem 2.1.3 $\Gamma^a_{b\,c}$ is commutative in its subscripts: $\boxed{\Gamma^a_{b\,c} = \Gamma^a_{c\,b}}$. (2.14)

Theorem 2.1.4 (Exercises 2.1.2 and 2.1.3) Suppose $\gamma$ is a geodesic curve that is affinely parameterized by *u,* and $\dot{x}^a \equiv \frac{dx^a}{du}$ are tangent vectors to $\gamma$. Then the lengths $L(u) \overset{(1.22)}{=} \sqrt{\left| g_{a\,b}\, \dot{x}^a \dot{x}^b \right|}$ of $\dot{x}^a$ are constant. If $\gamma$ is non-null, then there are constants *A* and *B*, $A \neq 0$, such that $u = As + B$.

Example 2.1.1 Determine which circles of latitude on Earth are geodesics.




Let a sphere have radius $a$. Using spherical coordinates $(r, \theta, \phi)$, the equation $\theta = \theta_0$ represents a circle of latitude, parallel to the $xy$-plane at a height of h = $a \cos\theta_0$. The radius of the circle is $\rho = a \sin\theta_0$. An arc in the circle has length $s = \rho\, \phi = a\, \phi \sin\theta_0$ where $\phi$ is the subtended angle. We can parameterize the circle with $u^1 = \theta = \theta_0$ and $u^2 = \phi = \dfrac{s}{a \sin\theta_0}$. This can be expressed as

$$u^A = \theta_0 \delta_1^A + \frac{s}{a \sin\theta_0} \delta_2^A \quad \text{for } A = 1, 2.$$

So,

$$\dot{u}^A \equiv \frac{du^A}{ds} = \frac{1}{a \sin\theta_0} \delta_2^A \quad \text{and } \ddot{u}^A = 0.$$

The geodesic equations are

$$\ddot{u}^A + \Gamma^A_{B\,C}\, \dot{u}^B\, \dot{u}^C = \frac{1}{a^2 \sin^2 \theta_0} \delta_2^B\, \delta_2^C\, \Gamma^A_{B\,C} = \frac{1}{a^2 \sin^2 \theta_0} \Gamma^A_{22} = 0 \quad \text{for } A = 1 \text{ and } 2$$

$\Leftrightarrow \quad \Gamma^A_{22} = 0$ for $A = 1$ and 2.

From Exercise 2.1.5, the only non-zero connection coefficients for the sphere are

$$\Gamma^1_{22} = \sin\theta \cos\theta \quad \text{and } \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta.$$

So, $\Gamma^2_{22} = 0$, and the equation $\Gamma^A_{22} = 0$ is only new information for $A = 1$, giving

$$\Gamma^1_{22} = \sin\theta_0 \cos\theta_0 = 0 \;\Leftrightarrow\; \theta_0 = \frac{\pi}{2} \text{ or } \theta_0 = 0.$$

Since $\theta_0 = 0$ is just a point, not a true latitude circle, the equator is the only latitude that is a geodesic. ■

Claim: The **covariant geodesic equations for a curve $\gamma$** are

$$g_{cb}\,\ddot{x}^b + \Gamma_{cab}\,\dot{x}^a\,\dot{x}^b = 0 \tag{2.1-4}$$

where $\dot{x}^a = \frac{dx^a}{du}$ and $u$ is an affine parameter for $\gamma$.

To confirm this, observe that the standard **contravariant geodesic equations** (2.12) for the curve $\gamma$ can be rewritten using $\dot{x}$ notation as

$$\ddot{x}^c + \Gamma^c_{ab}\,\dot{x}^a\,\dot{x}^b = 0\,. \tag{2.19}$$

To convert the covariant equation to this, we raise the 1st subscript of $\Gamma$:

$$0 = g^{dc}\,g_{cb}\,\ddot{x}^b + g^{dc}\,\Gamma_{cab}\,\dot{x}^a\,\dot{x}^b = \delta^d_b\,\ddot{x}^b + \Gamma^d_{ab}\,\dot{x}^a\,\dot{x}^b = \ddot{x}^d + \Gamma^d_{ab}\,\dot{x}^a\,\dot{x}^b$$
$$= \ddot{x}^c + \Gamma^c_{ab}\,\dot{x}^a\,\dot{x}^b \qquad ✓$$

The Lagrangian method is a way to solve geodesic equations (2.13/2.13) for a curve $\gamma$ without first solving for the connection coefficients. The Lagrangian can be thought of as the length of an infinitesimal portion of $\gamma$, except if $\gamma$ is a null curve then it cannot be integrated to get length of the curve.

Definition The **Lagrangian** is defined by

$$L = \frac{1}{2}\,g_{ab}\,\dot{x}^a\,\dot{x}^b \tag{2.1-5}$$

where $\dot{x}^a = \frac{dx^a}{du}$ and $u$ is an affine parameter for $\gamma$.

Since $g_{ab}$ is a function of $x^c$, $L$ is a function of both $x^c$ and $\dot{x}^c$, which we regard as $2N$ independent variables (because the magnitude and direction of a derivative has no correlation to the position vector's magnitude and direction). So, to be more clear, the book writes the Lagrangian as

$$L(\dot{x}^c, x^c) = \frac{1}{2}\,g_{ab}(x^c)\,\dot{x}^a\,\dot{x}^b.$$

Side Note When the Lagrangian is applied to spacetime, the eight independent variables include the four spacetime coordinates: $c\,t$, $x$, $y$, and $z$. That is, $x$, $y$, and $z$ are treated as independent from time, $t$. Rather, all the coordinates are dependent on the affine parameter, $\gamma$, for the path. In Euclidean physics we are used to treating position as a function of time, but this is not the viewpoint in spacetime.

Definition The **Euler-Lagrange equations for *L*** are

$$\boxed{\frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^c}\right) - \frac{\partial L}{\partial x^c} = 0}\,. \tag{2.17}$$

Theorem 2.1.5 The Euler-Lagrange equations for $L$ are equivalent to the covariant geodesic equations equations for $\gamma$.

Proof.

$L = \frac{1}{2} g_{ab}\, \dot{x}^a\, \dot{x}^b$.

$\frac{\partial L}{\partial \dot{x}^c} = g_{cb}\, \dot{x}^b$:

Since $g_{cb}$ is independent of $\dot{x}^c$,

$$\frac{\partial L}{\partial \dot{x}^c} = \frac{1}{2} g_{ab} \frac{\partial}{\partial \dot{x}^c} (\dot{x}^a\, \dot{x}^b) = \frac{1}{2} g_{ab} \left(\frac{\partial \dot{x}^a}{\partial \dot{x}^c}\, \dot{x}^b + \frac{\partial \dot{x}^b}{\partial \dot{x}^c}\, \dot{x}^a\right) = \frac{1}{2} g_{ab} (\delta^a_c\, \dot{x}^b + \delta^b_c\, \dot{x}^a)$$

$$= \frac{1}{2} g_{ab}\, \delta^a_c\, \dot{x}^b + \frac{1}{2} g_{ab}\, \delta^b_c\, \dot{x}^a = \frac{1}{2} g_{cb}\, \dot{x}^b + \frac{1}{2} g_{ac}\, \dot{x}^a = \frac{1}{2} g_{cb}\, \dot{x}^b + \frac{1}{2} g_{bc}\, \dot{x}^b \quad \checkmark$$

$$\frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^c}\right) = g_{cb} \frac{d\,\dot{x}^b}{du} + \frac{dg_{cb}}{du}\, \dot{x}^b = g_{cb}\, \ddot{x}^b + \dot{g}_{cb}\, \dot{x}^b\,.$$

$$\frac{\partial L}{\partial x^c} = \frac{1}{2} \partial_c g_{ab}\, \dot{x}^a\, \dot{x}^b\,.$$

Thus, equations (2.17) becomes

$$0 = \frac{d}{du}\left(\frac{\partial L}{\partial \dot{x}^c}\right) - \frac{\partial L}{\partial x^c} = g_{cb}\, \ddot{x}^b + \dot{g}_{cb}\, \dot{x}^b - \frac{1}{2} \partial_c g_{ab}\, \dot{x}^a\, \dot{x}^b\,.$$

Since

$$\dot{g}_{cb} = \frac{dg_{cb}}{du} = \frac{\partial g_{cb}}{\partial x^a} \frac{dx^a}{du} = \partial_a g_{cb}\, \dot{x}^a,$$

we get

$$0 = g_{cb}\, \ddot{x}^b + \partial_a g_{cb}\, \dot{x}^a\, \dot{x}^b - \frac{1}{2} \partial_c g_{ab}\, \dot{x}^a\, \dot{x}^b\,.$$

But,

$$\partial_a g_{c\,b}\,\dot{x}^a\,\dot{x}^b = \frac{1}{2}\,\partial_a g_{c\,b}\,\dot{x}^a\,\dot{x}^b + \frac{1}{2}\,\partial_a g_{c\,b}\,\dot{x}^a\,\dot{x}^b$$

$$= \frac{1}{2}\,\partial_a g_{c\,b}\,\dot{x}^a\,\dot{x}^b + \frac{1}{2}\,\partial_b g_{c\,a}\,\dot{x}^b\,\dot{x}^a = \frac{1}{2}\,(\partial_a g_{c\,b} + \partial_b g_{c\,a})\,\dot{x}^a\,\dot{x}^b\,,$$

yielding

$$0 = g_{c\,b}\,\ddot{x}^b + \frac{1}{2}\,(\partial_a g_{c\,b} + \partial_b g_{c\,a} - \partial_c g_{a\,b})\,\dot{x}^a\,\dot{x}^b \overset{(2.33)}{=} g_{c\,b}\,\ddot{x}^b + \Gamma_{c\,a\,b}\,\dot{x}^a\,\dot{x}^b \qquad \blacksquare$$

Definition  We call a coordinate $x^d$ **cyclic** or **ignorable** if $g_{a\,b}$ is not a function of $x^d$.

If $x^d$ is a cyclic coordinate, then $\frac{\partial L}{\partial x^d} = 0$ , and so $\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\partial L}{\partial \dot{x}^d}\right) = 0$. This implies that $\frac{\partial L}{\partial \dot{x}^d}$ is constant along $\gamma$. Define $p_d \equiv \dot{x}_d$ . It is proportional to velocity along the curve $\gamma$ and, hence, to momentum. So,

$$p_d = \dot{x}_d = g_{d\,b}\,\dot{x}^b \overset{(2.1.-5)}{=} \frac{\partial L}{\partial \dot{x}^d}$$

shows that the $d$-component of momentum is constant along $\gamma$. That is, the $d$-component of momentum is conserved if $x^d$ is cyclic. This gives access to immediate integrals of geodesic equations, as illustrated in the following example.

Example 2.1.2  The **Robertson-Walker line element** in spacetime (to be developed in Chapter 6) is defined by

$$g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \equiv \mathrm{d}t^2 - R(t)^2\,[\frac{1}{1-k\,r^2}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2 + r^2\,sin^2\theta\,\mathrm{d}\phi^2]$$

where $\mu$, $\nu$ = 0,1,2,3, $k$ is a constant, and $x^0 \equiv t$, $x^1 \equiv r$, $x^2 \equiv \theta$, $x^3 \equiv \phi$. Find the geodesic equations and use them to find the connection coefficients.

Note: Compare with the line element for a sphere: $\mathrm{d}s^2 = r^2\,\mathrm{d}\theta^2 + r^2\,\sin^2\theta\,\mathrm{d}\phi^2$.

Solution.  Letting dots denote differentiation with respect to an affine parameter $u$:

$$\dot{x}^0 = \dot{t},\;\; \dot{x}^1 = \dot{r},\;\; \dot{x}^2 = \dot{\theta},\;\; \dot{x}^3 = \dot{\phi}, \qquad \ddot{x}^0 = \ddot{t},\;\; \ddot{x}^1 = \ddot{r},\;\; \ddot{x}^2 = \ddot{\theta},\;\; \ddot{x}^3 = \ddot{\phi},$$

$$g_{00} = 1,\; g_{11} = -\frac{R^2}{1-k\,r^2},\; g_{22} = -R^2 r^2\,,\; g_{33} = -R^2 r^2 sin^2\theta,\;\; \text{all others are zero.}$$

The Lagrangian is

$$L \overset{(2.1-5)}{=} \frac{1}{2}\, g_{\mu\nu}\, \dot{x}^{\mu}\, \dot{x}^{\nu} = \frac{1}{2}[\dot{t}^2 - R^2(\frac{1}{1-k\,r^2}\dot{r}^2 + r^2\dot{\theta}^2 + r^2 sin^2\theta\, \dot{\phi}^2)]\,.$$

Denoting $R' = \frac{\mathrm{d}R}{\mathrm{d}t}$, the Euler-Lagrange equations (2.17), $\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\partial L}{\partial \dot{x}^c}\right) - \frac{\partial L}{\partial x^c} = 0$, can be computed:

$$\frac{\partial L}{\partial \dot{t}} = \dot{t}, \quad \frac{\partial L}{\partial \dot{r}} = -R^2 \frac{1}{1-k\,r^2}\, \dot{r}, \quad \frac{\partial L}{\partial \dot{\theta}} = -\,R^2 r^2 \dot{\theta}, \quad \frac{\partial L}{\partial \dot{\phi}} = -\,R^2 r^2 sin^2\theta\, \dot{\phi}\,,$$

Observe that $\frac{\mathrm{d}R}{\mathrm{d}u} = \frac{\mathrm{d}R}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}u} = R'\,\dot{t}$, so that

$$\frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial L}{\partial \dot{t}} = \ddot{t}, \quad \frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial L}{\partial \dot{r}} = -\,2\,R\,R'\,\frac{1}{1-k\,r^2}\,\dot{t}\,\dot{r} - R^2\,(\frac{1}{1-k\,r^2}\,\ddot{r} + \frac{2\,k\,r}{\left(1-k\,r^2\right)^2}\,\dot{r}^2\,)\,,$$

$$\frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial L}{\partial \dot{\theta}} = -\,2\,R\,R'\,r^2\,\dot{t}\,\dot{\theta} - r\,R^2(2\dot{r}\dot{\theta} + r\ddot{\theta})\,,$$

$$\frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial L}{\partial \dot{\phi}} = -\,2\,R\,R'\,r^2\,\sin^2\theta\,\dot{t}\,\dot{\phi} - R^2 r\,(2 sin^2\theta\,\dot{r}\dot{\phi} + 2r\,sin\theta\,cos\theta\,\dot{\theta}\,\dot{\phi} + r sin^2\theta\,\ddot{\phi})\,,$$

$$\frac{\partial L}{\partial t} = -\,R\,R'\,(\frac{1}{1-k\,r^2}\dot{r}^2 + r^2\dot{\theta}^2 + r^2 sin^2\theta\,\dot{\phi}^2)$$

since $\frac{\partial \dot{t}^2}{\partial t} = 2\dot{t}\,\frac{\partial \dot{t}}{\partial t} = 2\dot{t}\,\frac{\partial}{\partial t}\dot{t} = 2\dot{t}\,\frac{\partial}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}u} = 2\dot{t}\,\frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial t}{\partial t} = 2\dot{t}\,\frac{\mathrm{d}}{\mathrm{d}u}(1) = 0$

and $\frac{\partial r}{\partial t} = \frac{\partial \theta}{\partial t} = \frac{\partial \phi}{\partial t} = 0$,

$$\frac{\partial L}{\partial r} = -\,R^2\,(\frac{k\,r}{\left(1-k\,r^2\right)^2}\dot{r}^2 + r\dot{\theta}^2 + r sin^2\theta\,\dot{\phi}^2)\,,$$

$$\frac{\partial L}{\partial \theta} = -\,R^2\,r^2 sin\theta\,cos\theta\,\dot{\phi}^2$$

since $\frac{\partial \dot{\theta}^2}{\partial \theta} = 2\dot{\theta}\,\frac{\partial \dot{\theta}}{\partial \theta} = 2\dot{\theta}\,\frac{\partial}{\partial \theta}\dot{\theta} = 2\dot{\theta}\,\frac{\partial}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}u} = 2\dot{\theta}\,\frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial \theta}{\partial \theta} = 0$,

$$\frac{\partial L}{\partial \phi} = 0$$

since $\frac{\partial \dot{\phi}^2}{\partial \phi} = 2\dot{\phi}\,\frac{\partial \dot{\phi}}{\partial \phi} = 2\dot{\phi}\,\frac{\partial}{\partial \phi}\dot{\phi} = 2\dot{\phi}\,\frac{\partial}{\partial \phi}\frac{\mathrm{d}\phi}{\mathrm{d}u} = 2\dot{\phi}\,\frac{\mathrm{d}}{\mathrm{d}u}\frac{\partial \phi}{\partial \phi} = 0$.

This yields the Euler-Lagrange equations, $\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\partial L}{\partial \dot{x}^c}\right) - \frac{\partial L}{\partial x^c} = 0$:

$$(t):\quad \ddot{t} + R\,R'\left(\frac{1}{1-k\,r^2}\dot{r}^2 + r^2\dot{\theta}^2 + r^2 sin^2\theta\,\dot{\phi}^2\right) = 0,$$

$$(r):\quad -2\,R\,R'\,\frac{1}{1-k\,r^2}\,\dot{t}\,\dot{r} - R^2\left(\frac{1}{1-k\,r^2}\,\ddot{r} + \frac{2\,k\,r}{(1-k\,r^2)^2}\,\dot{r}^2\right)$$
$$+\ R^2\left(\frac{k\,r}{(1-k\,r^2)^2}\dot{r}^2 + r\dot{\theta}^2 + r sin^2\theta\,\dot{\phi}^2\right)$$
$$= -R^2\frac{1}{1-k\,r^2}\,\ddot{r} - 2\,R\,R'\,\frac{1}{1-k\,r^2}\,\dot{t}\,\dot{r}$$
$$- R^2\frac{k\,r}{(1-k\,r^2)^2}\dot{r}^2 + R^2 r\dot{\theta}^2 + R^2 r sin^2\theta\,\dot{\phi}^2 = 0$$

$$(\theta):\quad -R^2r^2\ddot{\theta} - 2\,R\,R'\,r^2\,\dot{t}\,\dot{\theta} - 2R^2\,r\,\dot{r}\,\dot{\theta} + R^2r^2 sin\theta\,cos\theta\,\dot{\phi}^2 = 0$$

$$(\phi):\quad -R^2r^2 sin^2\theta\,\ddot{\phi} - 2\,R\,R'\,r^2\,\sin^2\theta\,\dot{t}\,\dot{\phi}$$
$$-2R^2r\,sin^2\theta\,\dot{r}\,\dot{\phi} - 2R^2r^2\,sin\theta\,cos\theta\,\dot{\theta}\,\dot{\phi} = 0$$

or

$$(t):\quad \ddot{t} + R\,R'\left(\frac{1}{1-k\,r^2}\dot{r}^2 + r^2\dot{\theta}^2 + r^2 sin^2\theta\,\dot{\phi}^2\right) = 0,$$
$$(r):\quad \ddot{r} + \frac{2}{R}\,R'\,\dot{t}\,\dot{r} + \frac{k\,r}{1-k\,r^2}\dot{r}^2 - (1-k\,r^2)\,r\,\dot{\theta}^2 - (1-k\,r^2)\,r\,sin^2\theta\,\dot{\phi}^2 = 0$$
$$(\theta):\quad \ddot{\theta} + \frac{2}{R}\,R'\,\dot{t}\,\dot{\theta} + \frac{2}{r}\,\dot{r}\,\dot{\theta} - sin\theta\,cos\theta\,\dot{\phi}^2 = 0 \qquad (2.20)$$
$$(\phi):\quad \ddot{\phi} + R'\,\dot{t}\,\dot{\phi} + \frac{2}{r}\,\dot{r}\,\dot{\phi} + 2\,cot\theta\,\dot{\theta}\,\dot{\phi} = 0$$

This completes the geodesic equations. However, after-the-fact we can now easily find the connection coefficients by matching coefficients with geodesic equations (2.12):

$$\ddot{x}^a + \Gamma^a_{b\,c}\dot{x}^b\dot{x}^c = 0\ .$$

$$\ddot{t} = \ddot{x}^0 \Rightarrow \Gamma^0_{b\,c}:\quad \Gamma^0_{11}\,\dot{r}^2 + \Gamma^0_{22}\,\dot{\theta}^2 + \Gamma^0_{33}\,\dot{\phi}^2 = 0$$
$$\ddot{r} = \ddot{x}^1 \Rightarrow \Gamma^1_{b\,c}:\quad 2\,\Gamma^1_{01}\,\dot{t}\,\dot{r} + \Gamma^1_{11}\,\dot{r}^2 + \Gamma^1_{22}\,\dot{\theta}^2 + \Gamma^1_{33}\,\dot{\phi}^2 = 0$$
$$\ddot{\theta} = \ddot{x}^2 \Rightarrow \Gamma^2_{b\,c}:\quad 2\,\Gamma^2_{02}\,\dot{t}\,\dot{\theta} + 2\,\Gamma^2_{12}\,\dot{r}\,\dot{\theta} + \Gamma^2_{33}\,\dot{\phi}^2 = 0$$
$$\ddot{\phi} = \ddot{x}^3 \Rightarrow \Gamma^3_{b\,c}:\quad 2\,\Gamma^3_{03}\,\dot{t}\,\dot{\phi} + 2\,\Gamma^3_{13}\,\dot{r}\,\dot{\phi} + 2\,\Gamma^3_{23}\,\dot{\theta}\,\dot{\phi} = 0$$

$$\Rightarrow\ \Gamma^0_{11} = R\,R'\,\frac{1}{1-k\,r^2},\quad \Gamma^0_{22} = R\,R' r^2,\quad \Gamma^0_{33} = R\,R'\,r^2\,sin^2\theta$$

$$\Gamma^1_{01} = \Gamma^1_{10} = \frac{R'}{R},\quad \Gamma^1_{11} = \frac{k\,r}{1-k\,r^2},\quad \Gamma^1_{22} = -\,(1-k\,r^2)\,r,\quad \Gamma^1_{33} = -\,(1-k\,r^2)\,r\,sin^2\theta,$$

$$\Gamma^2_{02} = \Gamma^2_{20} = \frac{R'}{R},\quad \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r},\quad \Gamma^2_{33} = -\,sin\theta\,cos\theta,$$

$$\Gamma^3_{03} = \Gamma^3_{30} = \frac{R'}{R},\quad \Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r},\quad \Gamma^3_{23} = \Gamma^3_{32} = cot\theta. \qquad \blacksquare$$

Note that when $b \neq c$, we used $2\,\Gamma^a_{bc} = \Gamma^a_{bc} + \Gamma^a_{cb}$, spitting the equation (2.20) coefficient into two equal components.

Also note that there is no $\dot{\phi}^2$ in the last equation since $\frac{\partial L}{\partial \phi} = 0$. The parameter $\phi$ is cyclic, and so $\frac{\partial L}{\partial \dot{\phi}} = -\,R^2 r^2 sin^2\theta\,\dot{\phi} = -\,R^2 r^2 sin^2\theta\,\frac{d\phi}{du}$ is constant along the curve $\gamma$. The last geodesic equation is simply differentiation along $\gamma$ with respect to $u$, which means that this last equation can be integrated over the curve.

Finally, note that $x^0$ was specified as equal to $t$, not $c\,t$ as defined in Appendix A equation (A.3). This accounts for lack of “c” factors throughout, including no “c” factors in the connection coefficients.

An alternate derivation of equations (2.20) can be accomplished by computing the connection coefficients directly, which I did using Mathematica to perform the numerous operations required.

## 2.2 Parallel Vectors Along a Curve

### Euclidean Space

We begin in Euclidean space having a Cartesian coordinate system where points can be represented as Cartesian vectors $\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$, and with an alternate coordinate system $(u,v,w) = \left(u^1, u^2, u^3\right)$ where points can be represented as coordinate vectors $\boldsymbol{r} \overset{(1.16)}{=} u^i\,\mathbf{e}_i$ , where

$$\mathbf{e}_i \overset{(1.2-1)}{\equiv} \frac{\partial \boldsymbol{r}}{\partial u^i} = \frac{\partial x}{\partial u^i}\,\boldsymbol{i} + \frac{\partial y}{\partial u^i}\,\boldsymbol{j} + \frac{\partial z}{\partial u^i}\,\boldsymbol{k}.$$

Let $\gamma$ be a curve expressed parametrically as $u^i(t)$. With "dot" representing differentiation with respect to $t$,

$$\dot{\mathbf{e}}_i = \frac{\mathrm{d}\mathbf{e}_i}{\mathrm{d}t} = \frac{\partial \mathbf{e}_i}{\partial u^j}\frac{\mathrm{d}u^j}{\mathrm{d}t} = \dot{u}^j\,\partial_j \mathbf{e}_i \overset{(2.1-3)}{=} \dot{u}^j\,\Gamma^k_{i\,j}\,\mathbf{e}_k \quad \text{for } i, j, k = 1-3\,. \tag{2.2-1}$$

Definition  Let $\gamma$ be a curve represented parametrically by $u^i(t)$ and let $\mathrm{P}_0$ with parameter $t_0$ be a point on $\gamma$. Let $\boldsymbol{\lambda}_0$ be an $\mathbb{E}^3$-vector at $\mathrm{P}_0$ and $\boldsymbol{\lambda}(t) \equiv \boldsymbol{\lambda}_0$ be $\mathbb{E}^3$-vectors at points along $\gamma$. The vectors $\boldsymbol{\lambda}(t)$ have constant length and direction. We call this concept **parallel transport of the vector $\boldsymbol{\lambda}_0$ along the curve $\gamma$**, and the set of vectors $\{\boldsymbol{\lambda}(t)\}$ is called a **parallel field of vectors along $\gamma$ generated by parallel transport of $\boldsymbol{\lambda}_0$**.

Since there is no change of length or direction, the vectors $\lambda(t)$ satisfy the differential equation

$$\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}t} = \mathbf{0} \tag{2.21}$$

with initial condition $\boldsymbol{\lambda}(0) = \boldsymbol{\lambda}_0$.

Theorem 2.2.1  Let $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(t_0)$ be a vector and $\{\boldsymbol{\lambda}(t)\}$ a field of vectors along a curve $\gamma$ having constant length $|\boldsymbol{\lambda}_0|$. Then $\{\boldsymbol{\lambda}(t)\}$ is a parallel field of vectors iff $\boldsymbol{\lambda}(t)$ is the solution to the differential equations

$$\dot{\lambda}^i + \Gamma^i_{j\,k}\,\lambda^j\,\dot{u}^k = 0 \quad \text{for } i = 1-3, \tag{2.22}$$

having initial condition $\boldsymbol{\lambda}(t_0) = \boldsymbol{\lambda}_0$, where $\dot{\lambda}^i = \dfrac{\mathrm{d}\lambda^i}{\mathrm{d}t}$ and $j$ and $k$ range over $1-3$.

Proof.

$\{\boldsymbol{\lambda}(t)\}$ is a parallel field of vectors

$$\overset{(2.21)}{\Longleftrightarrow} \mathbf{0} = \dot{\boldsymbol{\lambda}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\lambda^i\, \mathbf{e}_i) = \dot{\lambda}^i\, \mathbf{e}_i + \lambda^i\, \dot{\mathbf{e}}_i \overset{(2.2\text{–}1)}{=} \dot{\lambda}^i\, \mathbf{e}_i + \lambda^i\, \dot{u}^j\, \Gamma^k_{ij}\, \mathbf{e}_k\,.$$

Exchanging subscripts in the 2nd term, $i \to j \to k \to i$, shows this is equivalent to

$$\dot{\lambda}^i\, \mathbf{e}_i + \lambda^j\, \dot{u}^k\, \Gamma^i_{jk}\, \mathbf{e}_i = 0,$$

which is equivalent to the three equations (2.22). ■

## N-Dimensional Riemannian and Pseudo Riemannian Spaces

For $N$-manifolds, we adopt the Euclidean result (2.22) as a basis-free definition. We postpone until Theorem 2.2.3 discussion of whether or not $\boldsymbol{\lambda}$ is a vector; i.e., whether this definition of a “vector” satisfies the vector coordinate transformation equation.

Definition We say that **a vector $\boldsymbol{\lambda}(u)$** is **generated by parallel transport of $\boldsymbol{\lambda}_0$ along the curve $\gamma$** if

$$\dot{\lambda}^a + \Gamma^a_{bc}\lambda^b\, \dot{x}^c = 0 \quad \text{for } a = 1 - N, \qquad (2.23)$$

where $b$ and $c$ range over $1 - N$, is satisfied along with initial condition $\boldsymbol{\lambda}_0 \equiv \boldsymbol{\lambda}(u_0)$. We call $\{\boldsymbol{\lambda}(u)\}$ a **parallel field of vectors along $\gamma$**. Note that $\Gamma^a_{bc}$ has changed from a scalar (in Euclidean space) to a scalar field (in the manifold). So, when taking derivatives, it is insightful to express definition (2.23) more explicitly as

$$\dot{\lambda}^a_{\mathrm{Q}} + (\Gamma^a_{bc})_{\mathrm{Q}}\, \lambda^b_{\mathrm{Q}}\, \dot{x}^c_{\mathrm{Q}} = 0 \quad \text{for } a = 1 - N, \qquad (2.23)$$

where $\gamma$ is parameterized with parameter $u$, points on $\gamma$ are $\mathrm{Q} = (\, x^a(u)\,)$, coordinates of points at Q are denoted $x^a_{\mathrm{Q}} = x^a(u)$, and the vector $\boldsymbol{\lambda}$ transported to point Q is $\lambda^a_{\mathrm{Q}} = \lambda^a(u)$.

The property of “parallel” vectors in Euclidean space that is of importance to geodesic theory is that the vectors maintain a constant angle with the $x$-, $y$-, and $z$-coordinate axes. In a manifold, we call vectors “parallel” if the vectors maintain a constant angle with the surface coordinate axes at each point. With this definition, if the manifold is a sphere in $\mathbb{E}^3$, vectors that are parallel on the sphere are not parallel in $\mathbb{E}^3$.

In this terminology, an affinely parameterized geodesic (definition 2.12) is a curve $\gamma = \gamma(u)$ in a manifold is characterized by the fact that its tangent vectors $\dot{x}^a$ form a parallel field of vectors along $\gamma$. Thus, by Exercises 2.1.2 and 2.1.3, the tangent vectors have the same length and, if $\gamma$ is non-null, $u = A\, s + B$. This suggests that parallel vectors $\boldsymbol{\lambda}(u)$ all have the same length, proved next.

Theorem 2.2.2 (Generalization of Exercise 2.1.2) If $\boldsymbol{\lambda}(u)$ is parallelly transported along $\gamma$ then the lengths $|\boldsymbol{\lambda}(u)|$ are constant.

Proof. Let $\dot{\lambda}^a$ denote $\frac{d\lambda^a}{du}$ and similarly for $\dot{L}$, $\dot{x}$, etc.

$\boldsymbol{\lambda}(u)$ is generated by parallel transport along $\gamma$

$$\Leftrightarrow\ \dot{\lambda}^a + \Gamma^a_{c\,d}\lambda^c\,\dot{x}^d = 0 \qquad \text{for } a = 1 - N, \tag{2.2-2}$$

$$\text{where}\ \ \Gamma^a_{c\,d} = \frac{1}{2}\, g^{ae}\,(\partial_c g_{ed} + \partial_d g_{ce} - \partial_e g_{cd}) \tag{2.2-3}$$

Let $L = |\boldsymbol{\lambda}(u)|$. Then $\pm L^2 = g_{ab}\lambda^a\lambda^b$. So

$$\pm 2\, L\, \dot{L} = \dot{g}_{ab}\lambda^a\lambda^b + g_{ab}(\lambda^a\dot{\lambda}^b + \lambda^b\dot{\lambda}^a) = \dot{g}_{ab}\lambda^a\lambda^b + 2\, g_{ab}\lambda^b\dot{\lambda}^a$$

$$\overset{(2.2-1)}{=}\ \dot{g}_{ab}\lambda^a\lambda^b - 2\, g_{ab}\lambda^b\,\Gamma^a_{c\,d}\,\lambda^c\,\dot{x}^d.$$

Since $\dot{g}_{ab} = (\partial_d g_{ab})\,\dot{x}^d$,

$$\pm 2\, L\, \dot{L} = \partial_d g_{ab}\lambda^a\lambda^b\dot{x}^d - 2\, g_{ab}\,\Gamma^a_{c\,d}\,\lambda^b\,\lambda^c\,\dot{x}^d$$

$$\overset{(2.2-11)}{=} \partial_d g_{ab}\lambda^a\lambda^b\dot{x}^d - g_{ab}\, g^{ae}\,(\partial_c g_{ed} + \partial_d g_{ce} - \partial_e g_{cd})\,\lambda^b\,\lambda^c\,\dot{x}^d$$

$$= \partial_d g_{ab}\lambda^a\lambda^b\dot{x}^d - \delta^e_b\,(\partial_c g_{ed} + \partial_d g_{ce} - \partial_e g_{cd})\,\lambda^b\,\lambda^c\,\dot{x}^d$$

$$= \partial_d g_{ab}\lambda^a\lambda^b\dot{x}^d - (\partial_c g_{bd} + \partial_d g_{cb} - \partial_b g_{cd})\,\lambda^b\,\lambda^c\,\dot{x}^d$$

$$= (\partial_d g_{ab}\lambda^a\lambda^b\dot{x}^d - \partial_d g_{cb}\lambda^b\,\lambda^c\,\dot{x}^d) - (\partial_c g_{bd}\lambda^b\,\lambda^c\,\dot{x}^d - \partial_b g_{cd}\,\lambda^b\,\lambda^c\,\dot{x}^d)$$

$$= (\,\partial_d g_{ab}\lambda^a\lambda^b\dot{x}^d \overset{(a\leftrightarrow c)}{-} \partial_d g_{ab}\lambda^b\,\lambda^a\,\dot{x}^d) - (\partial_c g_{bd}\lambda^b\,\lambda^c\,\dot{x}^d \overset{(b\leftrightarrow c)}{-} \partial_c g_{bd}\,\lambda^c\,\lambda^b\,\dot{x}^d)$$

$$= 0 - 0 = 0$$

$\Leftrightarrow\ \dot{L} = 0 \quad \Leftrightarrow \quad L$ is constant. ■

There is not necessarily a space encompassing a manifold in which “parallel” vectors point in the same direction. Rather, in a manifold, “parallel” transport of a vector along a curve $\gamma$ maintains a constant direction in relation to the tangent vector at each point of $\gamma$ as will be proven in the next section and illustrated in the following example.

The book works the next example for the simple case of transporting a vector $\boldsymbol{\lambda}$ around a latitude circle beginning and ending at a point P = $(\theta_0, 0)$. I work the more general case of transporting along a latitude arc from $P_0 = (\theta_0, \phi_0)$ to $P_1 = (\theta_0, \phi_1)$, and also along a longitude arc from $P_0 = (\theta_0, \phi_0)$ to $P_1 = (\theta_1, \phi_0)$, in preparation for Example 3.3.1. My example is quite long because I fill in the many, many details that are glossed over in the book.

Example 2.2.1 Let $\boldsymbol{\lambda}$ be a unit vector at a point P = $(\theta, \phi)$ on a sphere of radius $a$. Assume $\boldsymbol{\lambda}$ points east of south by an angle $\alpha$ as shown in Figure (a) below.

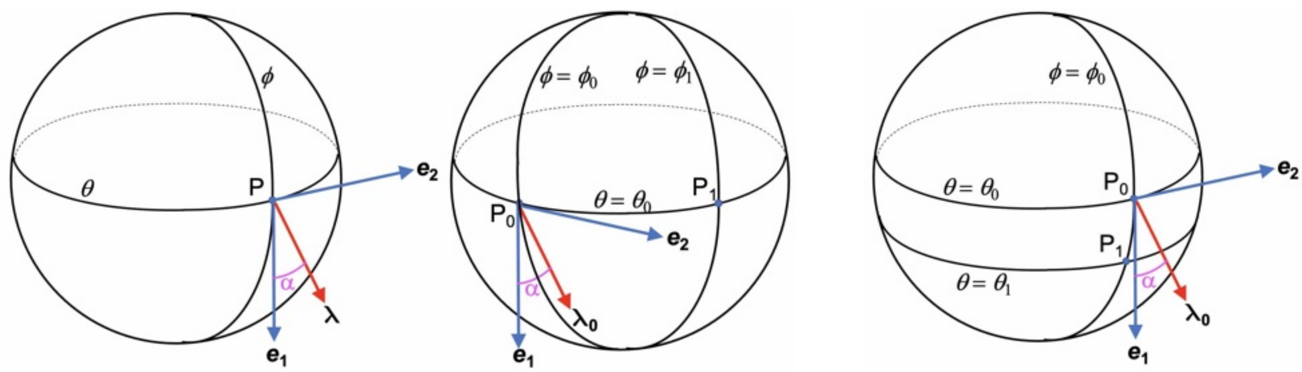



Figure a Figure b Figure c

#1. Using spherical coordinates, find the components with respect to the natural basis of the vector $\boldsymbol{\lambda}$ at P.

#2. Parameterize a latitude arc $\overset{\frown}{P_0 P_1}$ with an affine parameter $t$ and develop the parallel transport equations for $\boldsymbol{\lambda}$ (Figure b).

#3. Solve the parallel transport equations in #2 for transported vectors $\boldsymbol{\lambda}(t)$ and find the final vector at $P_1$.

#4. Show that the transported vectors $\boldsymbol{\lambda}(t)$ are unit vectors.

#5. (Exercise 2.2.3) Find the formula for the change in angle of $\boldsymbol{\lambda}(t)$ at a point $P_t$ on an arc and then show that only on the equator (a geodesic) do the vectors point in the same direction as well as maintain a constant angle with the tangent vectors. Show that a vector that is transported completely around a latitude circle that is near either the equator or the poles changes very little from the initial vector.

#6. Use the latitude arc to demonstrate that the result of vector transport on a curved surface can be path-dependent.

#7. (Not in book) Develop and solve the parallel transport equations for $\boldsymbol{\lambda}$ for an arc of longitude (Figure c). By a suitable rotation of the globe, the longitude arc becomes the equator, and so we know that all transported vectors make the same angle with the coordinate axes. Confirm this.

Solution. In spherical coordinates, the position vector for $\boldsymbol{\lambda}$ in 3-space is

$$\boldsymbol{r} = r \sin\theta \cos\phi\, \boldsymbol{i} + r \sin\theta \sin\phi\, \boldsymbol{j} + r \cos\theta\, \boldsymbol{k} \quad \text{for } 0 < \theta < \pi \text{ and } 0 \le \phi \le 2\pi.$$

A sphere of radius $a$ can be parameterized with parameters $(\theta,\phi)$. The natural basis at a point P on the sphere is

$$\mathbf{e}_1 = \mathbf{e}_\theta = \frac{\partial \boldsymbol{r}}{\partial \theta} = a \cos\theta \cos\phi\, \boldsymbol{i} + a \cos\theta \sin\phi\, \boldsymbol{j} - a \sin\theta\, \boldsymbol{k}$$

$$\mathbf{e}_2 = \mathbf{e}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = - a \sin\theta \sin\phi\, \boldsymbol{i} + a \sin\theta \cos\phi\, \boldsymbol{j}\,.$$

Observe that $\mathbf{e}_1$ and $\mathbf{e}_2$ are not unit vectors:

$$|\mathbf{e}_\theta| = a \quad \text{and} \quad |\mathbf{e}_\phi| = a \sin\theta \qquad (2.2\text{-}4)$$

In Exercise 2.1.5 it was shown that the metric tensor at a point P on the sphere can be represented in matrix form as

$$g_{AB} = \begin{pmatrix} a^2 & 0 \\ 0 & a^2 \sin^2\theta \end{pmatrix} \qquad (2.2\text{-}5)$$

and the only non-zero connection coefficients are

$$\Gamma^1_{22} = \sin\theta \cos\theta \text{ and } \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta. \qquad (2.2\text{-}6)$$

**#1** Let $\boldsymbol{S}_\theta$ be the unit vector at P pointing south (i.e., along $\mathbf{e}_1$), and $\boldsymbol{S}_\phi$ the unit vector at P pointing east (i.e., along $\mathbf{e}_2$) . Then, from (2.2-4) we have

$$\mathbf{e}_1 = a\, \boldsymbol{S}_\theta\,, \quad \mathbf{e}_2 = a \sin\theta\, \boldsymbol{S}_\phi\,,$$

and, so,

$$\boldsymbol{\lambda} = \cos\alpha\, \boldsymbol{S}_\theta + \sin\alpha\, \boldsymbol{S}_\phi = \frac{\cos\alpha}{a} \mathbf{e}_1 + \frac{\sin\alpha}{a \sin\theta} \mathbf{e}_2\,.$$

That is, we have shown that the vector $\boldsymbol{\lambda}$ has components

$$\boxed{\lambda^1 = \frac{\cos\alpha}{a} \text{ and } \lambda^2 = \frac{\sin\alpha}{a \sin\theta}}: \qquad (2.25)$$

**#2** Let $P_0 = (\theta_0, \phi_0)$ and $P_1 = P(\theta_0, \phi_1)$ be two points at the same latitude. We can parameterize the surface of the sphere using spherical coordinates:

$u^1 = \theta$ and $u^2 = \phi$.

We can then parameterize the latitude arc $\gamma = \overset{\frown}{P_0 P_1}$ with parameter *t* :

$u^1 = \theta = \theta_0$ and $u^2 = \phi = \phi_0 + t\,\Delta\phi \qquad$ for $0 \le t \le 1$

where $\quad \Delta\phi = \phi_1 - \phi_0$. (2.2-7)

Then $P_0 = P(0)$ and $P_1 = P(1)$. Also,

$$\dot{u}^1 = 0 \;\text{ and }\; \dot{u}^2 = \Delta\phi, \quad\text{or}\quad \dot{u}^A \equiv \frac{du^A}{dt} = \Delta\phi\,\delta_2^A.$$

In general $g_{AB}$, $g^{AB}$, and $\Gamma^A_{BC}$ are coordinate-dependent. However, we see from (2.2-5) and (2.2-6) that on $\gamma$, these quantities are the same at every point Q of $\gamma$. The parallel transport equations (2.22) for $\boldsymbol{\lambda}(t)$ on curve $\gamma$ become

$$\dot{\lambda}^1 + \Gamma^1_{BC}\,\lambda^B\,\dot{u}^C \overset{(2.2\text{–}5,\,2.2\text{–}6)}{=} \dot{\lambda}^1 + \Gamma^1_{22}\,\lambda^2\,\dot{u}^2 = \boxed{\dot{\lambda}^1 - \sin\theta_0\cos\theta_0\,\lambda^2\,\Delta\phi = 0}$$

$$\dot{\lambda}^2 + \Gamma^2_{BC}\,\lambda^B\,\dot{u}^C \overset{(2.2\text{–}5,\,2.2\text{–}6)}{=} \dot{\lambda}^2 + \Gamma^2_{12}\,\lambda^1\,\dot{u}^2 = \boxed{\dot{\lambda}^2 + \cot\theta_0\,\lambda^1\,\Delta\phi = 0} \quad ✓ \qquad (2.24)$$

**#3** Equation (2.25) yields an initial vector with components

$$\boxed{\lambda^1(0) = \frac{\cos\alpha}{a} \;\text{ and }\; \lambda^2(0) = \frac{\sin\alpha}{a\sin\theta_0}}\,. \qquad (2.2\text{-}8)$$

Set $\omega = \Delta\phi\,\cos\,\theta_0$ . (2.2-9)

We claim (Exercise 2.2.1) that the solution to the pair of DE's (2.24) that satisfies initial condition (2.2-8) is the vector having components

$$\boxed{\lambda^1(t) = \frac{\cos(\alpha - \omega\,t)}{a} \;\text{ and }\; \lambda^2(t) = \frac{\sin(\alpha - \omega\,t)}{a\sin\theta_0}}\;: \qquad (2.26)$$

Plugging $t = 0$ in equation (2.26) gives initial condition (2.2-8):

$$\lambda^1(0) \overset{(2.26)}{=} \frac{\cos\alpha}{a} \;\text{ and }\; \lambda^2(0) \overset{(2.26)}{=} \frac{\sin\alpha}{a\sin\theta_0} \quad ✓$$

The equations (2.24) yields

$$\dot{\lambda}^1 - \sin\theta_0 \cos\theta_0\, \lambda^2\, \Delta\phi \overset{(2.26)}{=} \frac{\omega \sin(\alpha-\omega t)}{a} - \sin\theta_0 \cos\theta_0 \frac{\sin(\alpha-\omega t)}{a \sin\theta_0}\, \Delta\phi$$

$$\overset{(2.2\text{-}9)}{=} \frac{\Delta\phi \cos\theta_0 \sin(\alpha-\omega t)}{a} - \frac{\cos\theta_0}{a} \sin(\alpha - \omega\, \mathrm{t})\, \Delta\phi = 0\,. \quad ✓$$

and

$$\dot{\lambda}^2 + \cot\theta_0\, \lambda^1\, \Delta\phi \overset{(2.26)}{=} -\frac{\omega \cos(\alpha-\omega t)}{a \sin\theta_0} + \frac{\cos\theta_0}{\sin\theta_0} \frac{\cos(\alpha-\omega t)}{a} \Delta\phi = 0 \quad ✓$$

Lastly, at point $P_1 = P(1)$, $t = 1$ and so equations (2.26) for the final vector become

$$\boxed{\lambda^1(1) = \frac{\cos(\alpha-\omega)}{a} \text{ and } \lambda^2(1) = \frac{\sin(\alpha-\omega)}{a \sin\theta_0}}\,. \quad ✓ \qquad (2.2\text{-}10)$$

**#4** All of the transported vectors on the latitude arc $\gamma$ are unit vectors:

$$|\boldsymbol{\lambda}(t)|^2 \overset{(1.80)}{=} g_{ab}\, \lambda^a(t)\, \lambda^b(t) = g_{11}\, [\lambda^1(t)]^2 + g_{22}\, [\lambda^2(t)]^2$$

$$= a^2 \frac{\cos^2(\alpha-\omega t)}{a^2} + a^2 \sin^2\theta_0 \frac{\sin^2(\alpha-\omega t)}{a^2 \sin^2\theta_0} = 1 \quad ✓$$

**#5** Define $\tilde{\boldsymbol{\lambda}}(t)$ as the unit vector at P($t$) that makes an angle of $\alpha$ with $\mathbf{e}_1(t) \equiv (\boldsymbol{e}_1)_{P_t}$. From equation (2.25), we have that

$$\tilde{\lambda}^1(t) = \frac{\cos\alpha}{a} \quad \text{and} \quad \tilde{\lambda}^2(t) = \frac{\sin\alpha}{a \sin\theta_0}\,. \qquad (2.2\text{-}11)$$

Let $\Delta(t)$ be the angle between $\boldsymbol{\lambda}(t)$ and $\tilde{\boldsymbol{\lambda}}(t)$. Our first task is to find the formula for $\Delta(t)$.

Because both vectors belong to the same coordinate system [i.e., at P($t$)], equation (1.81) can be used to compute the angle between them. Because they are unit vectors:

$$\cos\Delta(t) \overset{(1.81)}{=} g_{AB}\, \tilde{\lambda}^A(t)\, \lambda^B(t) = g_{11}\, \tilde{\lambda}^1(t)\, \lambda^1(t) + g_{22}\, \tilde{\lambda}^2(t)\, \lambda^2(t)$$

$$\overset{(2.2\text{-}5,\ 2.2\text{-}11,\ 2.26)}{=} a^2 \left[ \frac{\cos\alpha}{a} \frac{\cos(\alpha-\omega)}{a} + \sin^2\theta_0 \frac{\sin\alpha}{a \sin\theta_0} \frac{\sin(\alpha-\omega)}{a \sin\theta_0} \right]$$

$$= \cos(\alpha - \omega) \cos(\alpha) + \sin(\alpha - \omega) \sin(\alpha)$$

$$= \cos[(\alpha - \omega) - \alpha] = \cos\omega\,.$$

Hence, $\Delta(t) = \omega \overset{(2.2\text{-}9)}{=} \Delta\phi \cos\theta_0$. That is: $\boxed{\Delta(t) = \Delta\phi \cos\theta_0}$. ✓ (2.2-12)

Only at the equator do we have $\Delta(t) = \Delta\phi \, cos\frac{\pi}{2} = 0$ for all points on the latitude curve. So, the equator is the only latitude circle where all of the vectors point in the same direction. Thus, only at the equator do all the vectors maintain a constant angle with the tangent vectors $\boldsymbol{e}_2(t) \equiv (\boldsymbol{e}_2)_{P(t)}$. ✓

Near the equator, $\theta_0 \approx \frac{\pi}{2}$, so $\Delta(t) \approx 0$, which means that there is very little change in the angle of $\lambda(t)$. In particular, the final vector change of angle is $\lambda(2\pi)$, very small ✓

Near the poles, $\Delta(t) \approx \pm\, \Delta\phi$, so the final vector (where $\Delta\phi = 2\pi$) has $\Delta(1) \approx \pm 2\pi$, which is very little change in angle. ✓

While the result of transporting a vector completely around a latitude circle is approximately the same at the poles as at the equator, the interim transported vectors are quite different. Near the equator, all of the vectors are almost unchanged. At the poles, the transported vectors tilt more and more until they point 180° in the opposite direction, and then the tilt shrinks back to almost zero at the final vector.

**#6** From #5, $\Delta(1) = 0$ for an equatorial path $\gamma$ starting and ending at a point P, so the final and initial vectors point in the same direction. But, any other circular path through P can be considered to be a non-equatorial longitudinal path by a suitable rotation of the globe. Thus, $\Delta(1) \neq 0$ for any other circular path starting and ending at a point P and, hence, the final vector points in a different direction than the initial vector. This shows that the equation for the final vector at P is path-dependent. ✓

**#7** Let the arc $\gamma = \overset{\frown}{P_0 P_1}$ be the segment of the semi-circle $\phi = \phi_0$ from latitude $\theta_0$ to latitude $\theta_1$ as shown in Figure (c). As before, we assume the initial vector $\boldsymbol{\lambda}_0$, anchored at $P_0 = (\theta_0, \phi_0)$, points east of south by an angle $\alpha$ ; that is, it makes an angle $\alpha$ with basis vector $\boldsymbol{e}_1$.

We can parameterize $\gamma$ with parameter $t$ as follows:

$$u^1 = \theta = \theta_0 + t\,\Delta\theta \quad \text{for } t = 0 \text{ to } 1, \quad \text{where} \quad \Delta\theta = \theta_1 - \theta_0\,,$$
$$u^2 = \phi = \phi_0. \tag{2.2-13}$$

Then $\dot{u}^1 = \Delta\theta$ and $\dot{u}^2 = 0$. (2.2-14)

Unlike parallel transport along a latitude arc, the connection coefficients along this arc are coordinate-dependent; different at each point of $\overset{\frown}{P_0 P_1}$. $\Gamma^1_{22} = \sin\theta \cos\theta$ and $\Gamma^2_{21} = \cot\theta$ are functions of the arc parameter $t$. $(\Gamma^1_{22})_Q = \sin\theta(t) \cos\theta(t)$ and $(\Gamma^2_{21})_Q = \cot\theta(t)$ where $\theta(t) \overset{(2.2-13)}{=} \theta_0 + t\,\Delta\theta$.

Claim: the parallel transport equations (2.22) for the longitude arc $\gamma$ are

$$\boxed{\dot{\lambda}^1 = 0} \quad \text{and} \quad \boxed{\dot{\lambda}^2 + \lambda^2\,\Delta\theta \cot\theta = 0}: \qquad (2.2\text{-}15)$$

$$0 = \dot{\lambda}^1 + \Gamma^1_{BC}\,\lambda^B\,\dot{u}^C \overset{(2.2-6)}{=} \dot{\lambda}^1 + \Gamma^1_{22}\,\lambda^2\,\dot{u}^2 \overset{(2.2-14)}{=} \dot{\lambda}^1 + \Gamma^1_{22}\,\lambda^2\,(0) = \dot{\lambda}^1 \quad ✓$$

$$0 = \dot{\lambda}^2 + \Gamma^2_{BC}\,\lambda^B\,\dot{u}^C \overset{(2.2-6)}{=} \dot{\lambda}^2 + \Gamma^2_{12}\,\lambda^1\dot{u}^2 + \Gamma^2_{21}\,\lambda^2\,\dot{u}^1 \overset{(2.2-14)}{=} \dot{\lambda}^2 + \Gamma^2_{21}\,\lambda^2\,\Delta\theta$$

$$\overset{(2.2-6)}{=} \dot{\lambda}^2 + \lambda^2\,\Delta\theta \cot\theta \;✓$$

By equation (2.25), the initial vector $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(0)$ at the point $P_0$ has components

$$\boxed{\lambda^1(0) = \frac{\cos\alpha}{a} \text{ and } \lambda^2(0) = \frac{\sin\alpha}{a\sin\theta_0}}. \qquad (2.2\text{-}16)$$

At a point Q = P($t$) on the arc , let $\boldsymbol{\lambda}$ be the vector that makes the same angle $\alpha$ with $(\mathbf{e}_1)_Q$. Again, by equation (2.25), it has has components

$$\boxed{\lambda^1(t) = \frac{\cos\alpha}{\alpha}} \quad \text{and} \quad \boxed{\lambda^2(t) = \frac{\sin\alpha}{a\sin(\theta_0 + t\,\Delta\theta)}}. \qquad (2.2\text{-}17)$$

We claim that $\boldsymbol{\lambda} = \boldsymbol{\lambda}(t)$ satisfies the parallel transport equations (2.2-15) and has $\boldsymbol{\lambda}(0)$ as initial vector:

Initial vector: $\lambda^1(0) \overset{(2.2-17)}{=} \dfrac{\cos\alpha}{\alpha}$ ✓ $\qquad \lambda^2(0) \overset{(2.2-17)}{=} \dfrac{\sin\alpha}{a\sin\theta_0}$ ✓

$$\dot{\lambda}^1 \overset{(2.2-17)}{=} \frac{d}{dt}\frac{\cos\alpha}{\alpha} = 0 \;✓$$

$$\dot{\lambda}^2 = \frac{d}{dt}\frac{\sin\alpha}{a\sin(\theta_0 + t\,\Delta\theta)} = -\frac{\Delta\theta \sin\alpha \cos(\theta_0 + t\,\Delta\theta)}{a\sin^2(\theta_0 + t\,\Delta\theta)}$$

$$\dot{\lambda}^2 + \lambda^2\,\Delta\theta \cot\theta = -\frac{\Delta\theta \sin\alpha \cos(\theta_0 + t\,\Delta\theta)}{a\sin^2(\theta_0 + t\,\Delta\theta)} + \frac{\sin\alpha}{a\sin(\theta_0 + t\,\Delta\theta)}\,\Delta\theta\,\frac{\cos\theta}{\sin\theta}$$

$$\overset{(2.2-13)}{=} -\frac{\Delta\theta \sin\alpha \cos(\theta_0 + t\,\Delta\theta)}{a\sin^2(\theta_0 + t\,\Delta\theta)} + \frac{\sin\alpha}{a\sin(\theta_0 + t\,\Delta\theta)}\,\Delta\theta\,\frac{\cos(\theta_0 + t\,\Delta\theta)}{\sin(\theta_0 + t\,\Delta\theta)}$$

$$= 0 \;✓$$

We have shown that every transported vector makes the same angle with the coordinate axes. ✓

Thus, the final vector, $\boldsymbol{\lambda}_1$, also makes the same angle with the coordinate axes:

$$\boxed{\lambda^1(1) = \frac{\cos\alpha}{\alpha} \text{ and } \lambda^2(1) = \frac{\sin\alpha}{a\sin\theta_1}}. \qquad \blacksquare \qquad (2.2\text{-}18)$$

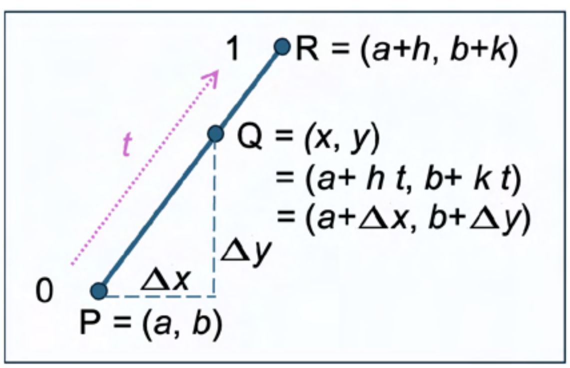
We have further learned that the great circles are the geodesics on a sphere because by using a suitable rotation they become the equator.

### Taylor Series of Functions of Several Variables

We take a short diversion to develop Taylor series in several variables expressed in index notation. This is important for generating first and second order approximations that we make when dealing with curvature. We use familiar calculus notation until the end, when we convert to index notation.

Let $f(x,y)$ have partials in an open region U containing points P = $(a,b)$ and R = $(a+h, b+k)$. We parameterize the line segment $\overline{\text{PR}}$ with $t$ :

$$\left.\begin{aligned} x &= a + t\,h \\ y &= b + t\,k \end{aligned}\right\} t \in [0, 1] \qquad (2.2\text{-}19)$$




$$\Rightarrow \quad \left.\begin{aligned} \frac{dx}{dt} &= \frac{d}{dt}(a + t\,h) = h \\ \frac{dy}{dt} &= \frac{d}{dt}(b + t\,k) = k \end{aligned}\right\} \qquad (2.2\text{-}20)$$

Let Q = $(x,y)$ be a point on the line segment. Define

$$\Delta x \equiv x - a \qquad \text{and} \qquad \Delta y = y - b \qquad (2.2\text{-}21)$$

Then

$$\Delta x \overset{(2.2-19)}{=} t\,h \qquad \text{and} \qquad \Delta y = t\,k \tag{2.2-22}$$

Define

$$F(t) \equiv f(a + t\,h,\, b + t\,k) \tag{2.2-23}$$

Then

$$F(t) \overset{(2.2-19)}{=} f(x,y) \quad \text{and} \qquad F(0) \overset{(2.2-23)}{=} f(a,b) \tag{2.2-24}$$

The standard Taylor series for $F$ is

$$F(t) = F(t_0) + F'(t_0)\,(t - t_0) + \frac{1}{2!}\,F''(t_0)\,(t - t_0)^2 + \ldots$$

In this case, $t_0 = 0$, so

$$F(\mathrm{t}) = F(0) + F'(0)\,\mathrm{t} + \frac{1}{2!}F''(0)\,t^2 + \ldots \tag{2.2-25}$$

$$\Rightarrow \quad F'(0) \;=\; \frac{\mathrm{d}F(0)}{\mathrm{d}t} = \frac{\partial F(0)}{\partial x}\,\frac{\mathrm{dx}}{\mathrm{d}t} + \frac{\partial F(0)}{\partial y}\,\frac{\mathrm{dy}}{\mathrm{d}t}$$

$$\overset{(2.2-24,\,2.2-20)}{=} \frac{\partial f(a,b)}{\partial x}\mathrm{h} + \frac{\partial f(a,b)}{\partial y}\mathrm{k} \tag{2.2-26}$$

$$\Rightarrow \quad F'(0)\,t \overset{(2-2.19)}{=} \frac{\partial f(a,b)}{\partial x}\Delta x + \frac{\partial f(a,b)}{\partial y}\Delta \mathrm{y} \tag{2.2-27}$$

and

$$F''(0) \;=\; \frac{\mathrm{d}F'(0)}{\mathrm{d}t} = \frac{\partial F'(0)}{\partial x}\,\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial F'(0)}{\partial y}\,\frac{\mathrm{d}y}{\mathrm{d}t}$$

$$\overset{(2.2-26,\,2.2-20)}{=} \frac{\partial^2 f(a,b)}{\partial x^2}h^2 + \frac{\partial f(a,b)}{\partial x\,\partial y}hk + \frac{\partial^2 f(a,b)}{\partial x\,\partial y}hk + \frac{\partial^2 f(a,b)}{\partial y^2}k^2$$

$$= \frac{\partial^2 f(a,b)}{\partial x^2}\mathrm{h} + 2\frac{\partial f(a,b)}{\partial x\,\partial y}\mathrm{k} + \frac{\partial^2 f(a,b)}{\partial y^2}\mathrm{k} \tag{2.2-28}$$

$$\Rightarrow \quad F''(0)\,t^2 \;=\; \frac{\partial^2 f(a,b)}{\partial x^2}h^2 t^2 + 2\frac{\partial f(a,b)}{\partial x\,\partial y}\mathrm{hk}t^2 + \frac{\partial^2 f(a,b)}{\partial y^2}k^2 t^2$$

$$= \frac{\partial^2 f(a,b)}{\partial x^2}\,(\Delta \mathrm{x})^2 + 2\frac{\partial f(a,b)}{\partial x\,\partial y}\,\Delta \mathrm{x}\,\Delta \mathrm{y} + \frac{\partial^2 f(a,b)}{\partial y^2}\,(\Delta \mathrm{y})^2 \tag{2.2-29}$$

Plugging equations (2.2-24), (2.2-27), and (2.2-29) for $F(0)$, $F'(0)\,t$, and $F''(0)\,t^2$, respectively, into equation (2.2-25) yields the Taylor series for a function of 2 variables:

$$f(x,y) = f(a,b) + \frac{\partial f(a,b)}{\partial x}\Delta x + \frac{\partial f(a,b)}{\partial y}\Delta y$$
$$+ \frac{1}{2!}\left[\frac{\partial^2 f(a,b)}{\partial x^2}(\Delta x)^2 + 2\frac{\partial f(a,b)}{\partial x\,\partial y}\Delta x\,\Delta y + \frac{\partial^2 f(a,b)}{\partial y^2}(\Delta y)^2\right] + \ldots \qquad (2.2\text{-}30)$$

Equation (2.2-30) easily extends to $f(x,y,z,\ldots,w)$. It is simpler to express if we use index notation:

$F(t) = f(x,y,z,\ldots,w)$ as $f(x^1, x^2, x^3, \ldots, x^N)$

$F(0) = f(a, b, c, \ldots, d) = f(x_0^1, x_0^2, x_0^3, \ldots, x_0^N)$

$\Delta x^i = x^i - x_0^i$ (for example, $\Delta x^2 = x^2 - x_0^2 = y - b = \Delta y$)

Then the Taylor series generalizes to

$$f(x^1, x^2, x^3, \ldots, x^N) = f(x_0^1, x_0^2, x_0^3, \ldots, x_0^N) + \sum_i \frac{\partial f\left(x_0^1, x_0^2, x_0^3, \ldots, x_0^N\right)}{\partial x^i}\Delta x^i$$
$$+ \frac{1}{2!}\left[\sum_i \sum_j \frac{\partial^2 f\left(x_0^1, x_0^2, x_0^3, \ldots, x_0^N\right)}{\partial x^i\,\partial x^j}\Delta x^i \Delta x^j\right] + \ldots \qquad (2.2\text{-}31)$$

In index notation, we write $f(x^a) \equiv f(x^1, x^2, x^3, \ldots, x^N)$, $f(x_0^a) \equiv f(x_0^1, x_0^2, x_0^3, \ldots, x_0^N)$, and $\xi^a \equiv \Delta x^a$, and this shortens to

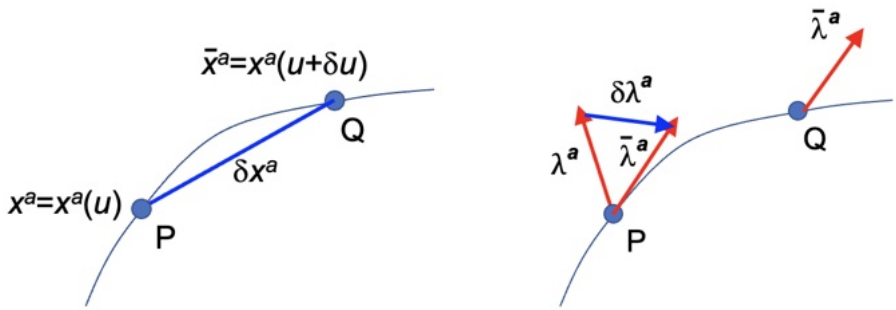
$$\boxed{f(x^a) = f(x_0^a) + \frac{\partial f(x_0^a)}{\partial x^b}\xi^b + \frac{1}{2!}\frac{\partial^2 f(x_0^a)}{\partial x^b\,\partial x^c}\xi^b\,\xi^c + \ldots} \qquad (2.2\text{-}32)$$

## Geodesic and Parallel Transport 1st Order Coordinate Independence in Manifolds

Definition The connection coefficients $\Gamma^a_{bc}$ are said to define **a connection on the manifold**. This enables an association between a vector in the tangent space at a point P with a parallel vector in the tangent space of a point Q.

Specifically, on a curve $\gamma$ in a manifold, if we label a vector at P by $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(u_0)$ then the parallel transport equation (2.23) along a geodesic curve $\gamma$ lets us solve for the parallel vector $\boldsymbol{\lambda}(u_1)$ at a point Q. The next theorem establishes that for points near P, $\boldsymbol{\lambda}(u_1)$ can be expressed as a vector in the coordinate system at P when approximated to the first order in the coordinate differences.

Theorem 2.2.3 In a differentiable manifold, let a curve $\gamma$ be parameterized by $u$, P = ( $x^a(u)$ ) = ($x^a$), Q = ( $x^a(u+\delta u)$ ) = ($\bar{x}^a$) , $\boldsymbol{\lambda} = (\lambda^a)$ a vector at P, and $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}^a)$ a parallel vector at Q. Then $\exists$ scalars $A^a_b \ni \bar{\lambda}^a = A^a_b\,\lambda^b + o(\delta u)^2$.




Proof. Denote $\dot{x}^a = \dfrac{dx^a}{du}$ and $\dot{\lambda}^a = \dfrac{d\lambda^a}{du}$. From the Taylor series expansion

$$\bar{x}^a \equiv x^a(u+\delta u) = x^a(u) + \dot{x}^a(u)\,\delta u + \tfrac{1}{2!}\,\ddot{x}^a(u)\,(\delta u)^2 + \ldots \tag{a}$$

we get the 1st order approximation (see figure)

$$\delta x^a \equiv \bar{x}^a - x^a \overset{(a)}{\approx} \dot{x}^a\,\delta \mathrm{u}. \tag{b}$$

From the Taylor series expansion

$$\bar{\lambda}^a \equiv \lambda^a(u+\delta u) = \lambda^a(u) + \dot{\lambda}^a(u)\,\delta u + \tfrac{1}{2!}\,\ddot{\lambda}^a(u)\,(\delta \mathrm{u})^2 + \ldots$$

we get the first order approximation

$$\delta\lambda^a \equiv \bar{\lambda}^a - \lambda^a \approx \dot{\lambda}^a\,\delta \mathrm{u}. \tag{c}$$

From parallel transport along the geodesic $\gamma$ we get that

$$\dot{\lambda}^a + \Gamma^a_{b\,c}\,\lambda^b\,\dot{x}^c \overset{(2.23)}{=} 0. \tag{d}$$

where $\Gamma^a_{b\,c} = \Gamma^a_{b\,c}(u)$ are scalars defined at P. So, $\bar{\lambda}^a$ has the first order approximation

$$\boxed{\bar{\lambda}^a \approx \lambda^a - \Gamma^a_{b\,c}\,\lambda^b\,\delta x^c}: \tag{2.27}$$

$$\bar{\lambda}^a \overset{(c)}{=} \lambda^a + \dot{\lambda}^a\,\delta\mathrm{u} \overset{(d)}{=} \lambda^a - \Gamma^a_{b\,c}\,\lambda^b\,\dot{x}^c\,\delta\mathrm{u} \overset{(b)}{=} \lambda^a - \Gamma^a_{b\,c}\,\lambda^b\,\delta x^c.$$

Define scalars $$A^a_b \equiv \delta^a_b - \Gamma^a_{b\,c}\,\delta x^c. \tag{2.28}$$

Since $\lambda^a = \lambda^b \delta^a_b$, we get the linear approximation

$$\bar{\lambda}^a \overset{(2.27)}{\approx} (\delta^a_b - \Gamma^a_{b\,c}\, \delta x^c)\, \lambda^b = A^a_b\, \lambda^b\,. \tag{2.2-33}$$

■

Corollary 2.2.3 A parallelly transported vector at a point Q on a curve can be approximated to the first order by a vector at a point P by using equation (2.27).

Definition Another name for the connection coefficients $\Gamma^a_{b\,c}$ is **Christoffel symbols of the first kind**, and other texts may denote them as $\{ {a \atop b\,c} \}$. A related quantity, $\Gamma_{a\,b\,c}$ or $[b\ c,\ a]$, is called **Christoffel symbols of the second kind**:

$$\boxed{\Gamma_{a\,b\,c} \equiv \frac{1}{2}\,(\partial_b g_{a\,c} + \partial_c g_{b\,a} - \partial_a g_{b\,c})} \tag{2.33}$$

Thus, $\Gamma_{d\,b\,c} \equiv \frac{1}{2}\,(\partial_b g_{d\,c} + \partial_c g_{b\,d} - \partial_d g_{b\,c})$ which, by equation (2.13), implies

$$\boxed{\Gamma^a_{b\,c} = g^{ad}\,\Gamma_{d\,b\,c}}\,. \tag{2.34}$$

Claim $\boxed{\Gamma_{a\,b\,c} = g_{a\,d}\,\Gamma^d_{b\,c}}$: (2.35)

$$g_{a\,d}\,\Gamma^d_{b\,c} \overset{(2.13)}{=} \frac{1}{2}\, g_{ad}\, g^{de}\,(\partial_b g_{ec} + \partial_c g_{be} - \partial_e g_{bc})$$
$$\overset{(1.78)}{=} \frac{1}{2}\,\delta^e_a\,(\partial_b g_{ec} + \partial_c g_{be} - \partial_e g_{bc}) = \frac{1}{2}\,(\partial_b g_{ca} + \partial_c g_{ab} - \partial_a g_{bc})$$
$$= \Gamma_{a\,b\,c} \qquad \checkmark$$

Equations (2.34) and (2.35) illustrate that $g$ raises and lowers indices of $\Gamma$. That is, they show that $\Gamma_{a\,b\,c}$ and $\Gamma^a_{b\,c}$ are *associated tensors*.

Observe that we can express the partials of the metric tensor in terms of the connection coefficients:

$$\partial_c\, g_{a\,b} \overset{(2.33)}{=} \Gamma_{a\,b\,c} + \Gamma_{b\,a\,c} \tag{2.36}$$

Notation Let $\mathfrak{g}$ denote the **metric tensor determinant** $|\, g_{a\,b}\,|$.

The book states without comment that

$$\partial_c \mathfrak{g} = \mathfrak{g}\; g^{ab}\; \partial_c g_{ab}.$$

This formula is a distraction to prove and is not discussed in the book. I have developed a proof using Levi-Civita symbology that I have labeled "Exercise 2.2.7". I saved it in the folder with the other exercises I have worked.

Combining this equation with equation (2.36), the book goes on to derive

$$\partial_c g \overset{(2.36)}{=} g\, g^{ab} \left( \Gamma_{a\,b\,c} + \Gamma_{b\,a\,c} \right) = g \left( \Gamma^b_{b\,c} + \Gamma^a_{a\,c} \right) = 2\, g\, \Gamma^a_{a\,c}$$

which implies

$$\Gamma^a_{a\,b} = \frac{1}{2}\, g^{-1} \partial_b g = \frac{1}{2}\, \partial_b( \ln |g| ) \tag{2.37}$$

since $\frac{d}{dx} \ln u = \frac{1}{u} \frac{du}{dx}$. Alternate expressions are

$$\Gamma^a_{a\,b} = \partial_b \ln |g|^{\frac{1}{2}} \quad \text{and} \quad \Gamma^a_{a\,b} = |g|^{-\frac{1}{2}}\, \partial_b |g|^{\frac{1}{2}} \tag{2.38}$$

since $\frac{d}{dx} \ln x^{\frac{1}{2}} = \frac{1}{2\,x} = x^{-\frac{1}{2}} \frac{d}{dx} x^{\frac{1}{2}}$.

Definition $\boldsymbol{X^{a}_{b'\,c'}} \equiv \partial_{b'} X^a_{c'} = \dfrac{\partial^2 x^a}{\partial x^{b'}\, \partial x^{c'}}$ .

The following identity is useful in Exercise 2.25.

Theorem 2.2.4 $X^{a'}_{b'\,c} = X^d_{b'} X^{a'}_{c\,d}$

Proof. $X^{a'}_{b'\,c} = \dfrac{\partial X^{a'}_c}{\partial x^{b'}} = \dfrac{\partial X^{a'}_c}{\partial x^d} \dfrac{\partial x^d}{\partial x^{b'}} = X^d_{b'} X^{a'}_{c\,d}$ ■

Theorem 2.2.5 $\Gamma_{d\,e\,c}$ transforms as

$$\boxed{\Gamma_{d'\,e'\,c'} = X^b_{d'} X^f_{e'} X^a_{c'} \Gamma_{b\,f\,a} + X^a_{e'\,c'} X^b_{d'}\, g_{a\,b}} \tag{2.2-34}$$

Proof. First,

$$\Gamma_{b\,f\,a} \overset{(2.33)}{=} \frac{1}{2} (\partial_f\, g_{b\,a} + \partial_a\, g_{f\,b} - \partial_b\, g_{f\,a}). \tag{a}$$

Next, $g_{a\,b}$ is a (metric) tensor, so it transforms as

$$g_{c'\,d'} = X^a_{c'} X^b_{d'}\, g_{a\,b}\,. \tag{b}$$

Thus,

$$\partial_{e'} g_{c' d'} \overset{(b)}{=} (\partial_{e'} X^a_{c'}) X^b_{d'} g_{ab} + X^a_{c'} (\partial_{e'} X^b_{d'}) g_{ab} + X^a_{c'} X^b_{d'} (\partial_{e'} g_{ab})$$
$$= (X^a_{e' c'} X^b_{d'} + X^a_{c'} X^b_{e' d'} + X^a_{c'} X^b_{d'} \partial_{e'}) g_{ab} . \qquad \text{(c)}$$

But,

$$\partial_{e'} = \frac{\partial}{\partial x^{e'}} = \frac{\partial}{\partial x^f} \frac{\partial x^f}{\partial x^{e'}} = X^f_{e'} \partial_f. \qquad \text{(d)}$$

So,

$$\partial_{e'} g_{c' d'} \overset{(c,\, d)}{=} (X^a_{e' c'} X^b_{d'} + X^a_{c'} X^b_{e' d'} + X^a_{c'} X^b_{d'} X^f_{e'} \partial_f) g_{ab} . \qquad \text{(e)}$$

By exchanging $e' \to c' \to d' \to e'$ and $a \to b \to f \to a$ we also get

$$\partial_{c'} g_{d' e'} = (X^b_{c' d'} X^f_{e'} + X^b_{d'} X^f_{c' e'} + X^b_{d'} X^f_{e'} X^a_{c'} \partial_a) g_{bf} . \qquad \text{(f)}$$

Making yet another exchange, $c' \to d' \to e' \to c'$ and $b \to f \to a \to b$ yields

$$\partial_{d'} g_{e' c'} = (X^f_{d' e'} X^a_{c'} + X^f_{e'} X^a_{d' c'} + X^f_{e'} X^a_{c'} X^b_{d'} \partial_b) g_{fa} \qquad \text{(g)}$$

Hence,

$$\Gamma_{d' e' c'} \overset{(2.33)}{=} \frac{1}{2} (\partial_{e'} g_{d' c'} + \partial_{c'} g_{e' d'} - \partial_{d'} g_{e' c'})$$

$$\begin{aligned}
&\overset{(e,\, f,\, g)}{=} \frac{1}{2} (X^a_{e' c'} X^b_{d'} + X^a_{c'} X^b_{e' d'} + X^b_{d'} X^f_{e'} X^a_{c'} \partial_f) g_{ab} \\
&+ \quad \frac{1}{2} (X^b_{c' d'} X^f_{e'} + X^b_{d'} X^f_{c' e'} + X^b_{d'} X^f_{e'} X^a_{c'} \partial_a) g_{bf} \\
&- \quad \frac{1}{2} (X^f_{d' e'} X^a_{c'} + X^f_{e'} X^a_{d' c'} + X^f_{e'} X^a_{c'} X^b_{d'} \partial_b) g_{fa}
\end{aligned}$$

$$\begin{aligned}
&= X^b_{d'} X^f_{e'} X^a_{c'} \frac{1}{2} (\partial_f g_{ab} + \partial_a g_{fb} - \partial_b g_{fa}) \\
&+ \frac{1}{2} (X^a_{e' c'} X^b_{d'} g_{ab} + X^b_{d'} X^f_{c' e'} g_{bf}) \\
&+ \frac{1}{2} (X^a_{c'} X^b_{d' e'} g_{ab} - X^f_{d' e'} X^a_{c'} g_{fa}) \\
&+ \frac{1}{2} (X^b_{c' d'} X^f_{e'} g_{bf} - X^f_{e'} X^a_{d' c'} g_{fa})
\end{aligned}$$

$$\overset{(a)}{=} X^b_{d'} X^f_{e'} X^a_{c'} \Gamma_{bfa} + \frac{1}{2}\left(X^a_{e'\,c'} X^b_{d'} g_{ab} \overset{(f\to a)}{+} X^a_{c'\,e'} X^b_{d'} g_{ba}\right)$$
$$+ \frac{1}{2}\left(X^a_{c'} X^b_{d'\,e'} g_{ab} \overset{(f\to b)}{-} X^a_{c'} X^b_{d'\,e'} g_{ba}\right)$$
$$+ \frac{1}{2}\left(X^b_{c'\,d'} X^f_{e'} g_{bf} \overset{(a\to b)}{-} X^b_{d'\,c'} X^f_{e'} g_{fb}\right)$$

$$= X^b_{d'} X^f_{e'} X^a_{c'} \Gamma_{bfa} + X^a_{e'\,c'} X^b_{d'} g_{ab} \quad \blacksquare$$

Theorem 2.2.6 (Exercise 2.2.4) $\Gamma^a_{bc}$ transforms as

$$\boxed{\Gamma^{a'}_{b'\,c'} = \Gamma^d_{fg} X^{a'}_d X^f_{b'} X^g_{c'} + X^d_{c'\,b'} X^{a'}_d} \tag{2.32}$$

Proof.

$$\Gamma^{h'}_{e'\,c'} \overset{(2.34)}{=} \Gamma_{d'\,e'\,c'}\, g^{h'\,d'} = \Gamma_{d'\,e'\,c'}\, X^{h'}_j X^{d'}_i\, g^{ij}$$
$$\overset{(2.2-34)}{=} \left(X^b_{d'} X^f_{e'} X^a_{c'} \Gamma_{bfa} + X^a_{e'\,c'} X^b_{d'} g_{ab}\right) X^{h'}_j X^{d'}_i\, g^{ij}$$
$$= \left(X^b_{d'} X^{d'}_i\right) X^f_{e'} X^a_{c'} X^{h'}_j g^{ij} \Gamma_{bfa} + X^a_{c'\,e'} \left(X^b_{d'} X^{d'}_i\right) X^{h'}_j\, g^{ij} g_{ab}$$
$$= \delta^b_i X^f_{e'} X^a_{c'} X^{h'}_j\, g^{ij} \Gamma_{bfa} + \delta^b_i X^a_{c'\,e'} X^{h'}_j\, g^{ij} g_{ab}$$
$$= X^f_{e'} X^a_{c'} X^{h'}_j \left(g^{bj} \Gamma_{bfa}\right) + X^a_{c'\,e'} X^{h'}_j \left(g^{bj} g_{ab}\right)$$
$$= X^f_{e'} X^a_{c'} X^{h'}_j \Gamma^j_{fa} + X^a_{c'\,e'} X^{h'}_j \delta^j_a$$
$$= X^f_{e'} X^a_{c'} X^{h'}_j \Gamma^j_{fa} + X^j_{c'\,e'} X^{h'}_j$$

Replacing $h' \to a'$, $e' \to b'$, $j \to d$, and $a \to g$ results in

$$\Gamma^{a'}_{b'\,c'} = X^f_{b'} X^g_{c'} X^{a'}_d \Gamma^d_{fg} + X^d_{c'\,b'} X^{a'}_d \quad \blacksquare$$

Theorem 2.2.7 Parallel transport on a curve, as defined by equation (2.23), is coordinate-independent to a first order approximation in coordinate differences.

Proof. Let a curve, $\gamma$, be endowed with primed and unprimed coordinate systems at each of its points. Let $\boldsymbol{\lambda}$, a vector in the tangent plane at P, be denoted by $\lambda^a$ in the unprimed coordinate system at P = $(x^a)$ and by $\lambda^{a'}$ in the primed coordinate system at P = $(x^{a'})$. Let $\overline{\lambda}^a$ and $\overline{\lambda}^{a'}$ denote objects in the unprimed and primed coordinate systems, respectively, at Q. That is, the vector $\boldsymbol{\lambda}$ is parallely transported to an object $\overline{\lambda}^a$ at the point Q = $(\overline{x}^a)$, and to a possibly different object $\overline{\lambda}^{a'}$ at Q = $(\overline{x}^{a'})$ .

To prove that $\bar{\lambda}^a$ is coordinate-independent (i.e., a vector at Q with respect to the coordinate system at P), we must show that the coordinate transformation equation (1.70) holds at Q: $\bar{\lambda}^a (X_a^{e'})_Q = \bar{\lambda}^{e'}$. Equivalently, if we treat the primed system as the primary one, we must show

$$\bar{\lambda}^{a'} \left(X_{a'}^e\right)_Q = \bar{\lambda}^e \tag{2.2-35}$$

where $\left(X_{a'}^e\right)_Q$ denotes $\frac{\partial x^e}{\partial x^{a'}}$ evaluated at Q: $\left(X_{a'}^e\right)_Q = \frac{\partial \bar{x}^e}{\partial \bar{x}^{a'}}$. The partials $\frac{\partial x^e}{\partial x^{a'}}$ and $\frac{\partial \bar{x}^e}{\partial \bar{x}^{a'}}$ are taken in the primed coordinate systems at P and Q, respectively.

We saw in equation (2.27) that the object $\bar{\lambda}^e$ has the first order approximation

$$\bar{\lambda}^e \approx \lambda^e - \Gamma^e_{f\,g}\, \lambda^f\, \delta x^g \tag{2.2-36}$$

where

$$\delta x^g \equiv \bar{x}^g - x^g \text{ are unprimed coordinate differences between P and Q} \tag{2.2-37}$$

and, hence, the object $\bar{\lambda}^{a'}$ has the first order expression

$$\bar{\lambda}^{a'} \approx \lambda^{a'} - \Gamma^{a'}_{b'\,c'}\, \lambda^{b'}\, \delta x^{c'} \tag{2.2-38}$$

where

$$\delta x^{c'} \equiv \bar{x}^{c'} - x^{c'} \text{ are primed coordinate differences between P and Q.} \tag{2.2-39}$$

We seek a first order expression for the remaining term, $\left(X_{a'}^e\right)_Q$. We begin by developing Taylor series in the primed coordinate system at P, first for $\bar{x}^e$ and then for $\left(X_{a'}^e\right)_Q$. Denote a general function of multiple variables in the coordinate system at Q by $f(\bar{x}^e) \equiv f(\bar{x}^1, \ldots, \bar{x}^N)$, where each unprimed parameter $\bar{x}^e = \bar{x}^e(\bar{x}^{1'}, \ldots, \bar{x}^{N'})$ is a function of primed parameters. The general Taylor series for this function of multiple variables was developed in equation (2.2-32). Applying it to $f(\bar{x}^e)$ yields

$$f(\bar{x}^e) = f(x^e) + \frac{\partial f(x^e)}{\partial x^{f'}} \delta x^{f'} + \frac{1}{2!} \frac{\partial^2 f(x^e)}{\partial x^{g'}\, \partial x^{f'}} \delta x^{f'} \partial x^{g'} + \ldots$$

Set $f(x^e) = x^e$. Then $f(\bar{x}^e) = \bar{x}^e$, and

$$\begin{aligned} \bar{x}^e &= x^e + \frac{\partial x^e}{\partial x^{f'}} \delta x^{f'} + \frac{1}{2!} \frac{\partial^2 x^e}{\partial x^{g'}\, \partial x^{f'}} \delta x^{f'} \partial x^{g'} + \ldots \\ &= x^e + X^e_{f'}\, \delta x^{f'} + \frac{1}{2!} X^e_{g'\,f'} \delta x^{f'} \partial x^{g'} + \ldots \end{aligned} \tag{2.2-40}$$

Taking the partial derivative with respect to $x^{a'}$ (in the primed coordinate system at P) yields

$$\left(X^e_{a'}\right)_Q = \frac{\partial \overline{x}^e}{\partial x^{a'}} = \frac{\partial x^e}{\partial x^{a'}} + \frac{\partial^2 x^e}{\partial x^{a'}\,\partial x^{f'}}\,\delta x^{f'} + \frac{1}{2!}\,\frac{\partial^3 x^e}{\partial x^{a'}\,\partial x^{g'}\,\partial x^{f'}}\,\delta x^{f'}\partial x^{g'} + \ldots$$

$$= X^e_{a'} + X^e_{a'\,f'}\,\delta x^{f'} + \frac{1}{2!}\,X^e_{a'\,g'\,f'}\delta x^{f'}\partial x^{g'} + \ldots \qquad (2.2\text{-}41)$$

since $\frac{\partial}{\partial x^{a'}}(\delta x^{f'}) \overset{(2.2-39)}{=} \frac{\partial}{\partial x^{a'}}(\overline{x}^{f'} - x^{f'}) = \delta^{f'}_{a'} - \delta^{f'}_{a'} = 0.$

From (2.2-41), we get a first order approximation for $\left(X^e_{a'}\right)_Q$ :

$$\left(X^e_{a'}\right)_Q = X^e_{a'} + X^e_{d'\,a'}\,\delta x^{d'}. \qquad (2.2\text{-}42)$$

Substituting equations (2.2-36, 2.2-38, and 2.2-42) into equation (2.2-35) reduces the problem to that of showing

$$(\lambda^{a'} - \Gamma^{a'}_{b'\,c'}\,\lambda^{b'}\,\delta x^{c'})\,(X^e_{a'} + X^e_{d'\,a'}\,\delta x^{d'}) = \lambda^e - \Gamma^e_{f\,g}\,\lambda^f\,\delta x^g\,.$$

Since, $\lambda^{a'} X^e_{a'} \overset{(1.70)}{=} \lambda^e$, when we multiply out LHS, we get

$$\lambda^{a'} X^e_{d'\,a'}\,\delta x^{d'} - \Gamma^{a'}_{b'\,c'}\,\lambda^{b'}[X^e_{a'}\,\delta x^{c'} + X^e_{d'\,a'}\,\delta x^{c'}\delta x^{d'}] = -\Gamma^e_{f\,g}\,\lambda^f\,\delta x^g$$

Setting the second order term $\delta x^{c'}\delta x^{d'}$ to zero completes the elimination of all 2nd order and higher terms and reduces this to the 1st order approximation

$$X^e_{d'\,a'}\,\lambda^{a'}\,\delta x^{d'} - \Gamma^{a'}_{b'\,c'}\,X^e_{a'}\,\lambda^{b'}\,\delta x^{c'} = -\Gamma^e_{f\,g}\,\lambda^f\,\delta x^g\,.$$

Since $\lambda^a$ is a vector, $\lambda^f \overset{(1.70)}{=} X^f_{b'}\,\lambda^{b'}$, and the problem is now reduced to showing that

$$\Gamma^{a'}_{b'\,c'}\,(X^e_{a'}\,\lambda^{b'}\,\delta x^{c'}) - \Gamma^e_{f\,g}\,X^f_{b'}\,\lambda^{b'}\delta x^g - X^e_{d'\,a'}\,\lambda^{a'}\delta x^{d'} = 0.$$

We wish to factor $X^e_{a'}\,\lambda^{b'}\,\delta x^{c'}$ out of all three terms on LHS. The first term is ready. In the second term:

$$\Gamma^e_{f\,g} = \Gamma^d_{f\,g}\,\delta^{\,e}_{\,d} \overset{(1.68)}{=} \Gamma^d_{f\,g}\,X^{a'}_d\,X^e_{a'} \quad \text{and}$$

$$\delta x^g \overset{(2.2-37)}{=} \overline{x}^g - x^g \overset{(2.2-40)}{=} X^g_{c'}\,\delta x^{c'} \text{ to the first order.}$$

So the second term equals $-\Gamma^d_{fg} X^{a'}_d X^f_{b'} X^g_{c'} (X^e_{a'} \lambda^{b'} \delta x^{c'})$.

For the third term:

Changing $a' \to b'$ and $d' \to c'$ yields $X^e_{c'b'} \lambda^{b'} \delta x^{c'}$.

Also, $X^e_{c'b'} = X^d_{c'b'} \delta^e_d \overset{(1.68)}{=} X^d_{c'b'} X^{a'}_d X^e_{a'}$.

So, the third term equals $X^d_{c'b'} X^{a'}_d (X^e_{a'} \lambda^{b'} \delta x^{c'})$

The problem is now reduced to showing that

$$\Gamma^{a'}_{b'c'} - \Gamma^d_{fg} X^{a'}_d X^f_{b'} X^g_{c'} - X^d_{c'b'} X^{a'}_d = 0, \tag{2.31}$$

or

$$\Gamma^{a'}_{b'c'} = \Gamma^d_{fg} X^{a'}_d X^f_{b'} X^g_{c'} + X^d_{c'b'} X^{a'}_d,$$

which is equation (2.32), shown in Theorem 2.2.6 to be how $\Gamma^a_{bc}$ transforms. ■

Corollary 2.2.7 The geodesic definition (2.12–2.13) is coordinate-independent to a first order approximation.

Proof. By Theorem 2.27, parallel transport of a vector $\lambda^a$ at a point P = $(x^a)$ to a vector $\bar{\lambda}^a$ at a nearby point Q = $(x^a + \delta x^a)$ is approximately coordinate-independent. Since we can express the geodesic definition (2.12) in terms of parallel transport (2.23) by setting $\lambda^a = \frac{dx^a}{du}$, it follows that the geodesic definition is coordinate-independent to a first order approximation.

Observation Being approximately coordinate-independent is a sufficient condition to ensure generation of a unique geodesic path. Consider starting at P and traveling in a direction determined by a vector $\boldsymbol{\lambda}$ for a small distance $\delta x^a$ to a point $P_1$. Chose a coordinate system at random at $P_1$ and travel in the direction of the transported vector for another distance $\delta x^a$ to a point $P_2$. Stop the process after, say, 10 steps. Now cut $\delta x^a$ by a factor of 10 and repeat the process for 100 steps; then another factor of 10 and 1000 steps; etc. Requiring coordinate-independence to first order forces the transported vectors to increasingly approach true coordinate-independence as the step distance decreases, converging in the limit to a single geodesic path with true coordinate-independence for $\boldsymbol{\lambda}$ at every point.

## 2.3 Absolute and covariant differentiation

In this section we develop tensor derivatives. We assume henceforth that vectors are analytic, and in particular that the order of differentiation does not matter. Partial and total differentiation as defined in Euclidean space do not return tensors. Thus, we are led to define two new concepts, "absolute derivative" along a curve and "covariant derivative" for a region of a manifold or an entire manifold.

### Absolute Derivative

We begin with a curve $\gamma$, parameterized by $u$, in a manifold M. Points of M have the form P = $(x^a)$ and points of $\gamma$ can be expressed P = $(x^a(u))$.

Definition The **total derivative of vector $\boldsymbol{\lambda^a}$** is

$$\boxed{\frac{d\lambda^a}{du} = \lim_{\delta u \to 0} \frac{\lambda^a(u+\delta u) - \lambda^a(u)}{\delta u}}, \tag{2.41}$$

assuming the limit exists.

Theorem 2.3.1 $\frac{d\lambda^a}{du}$ is not a vector.

Proof. In order to be a vector, $\frac{d\lambda^a}{du}$ would transform as $\frac{d\lambda^{a'}}{du} = X^{a'}_{b} \frac{d\lambda^b}{du}$.

However, since $\lambda^a$ is a vector it transforms as $\lambda^{a'} = X^{a'}_{b}\, \lambda^b$. Because

$$\frac{d}{du} X^{a'}_{b} = \frac{d}{du} \frac{\partial x^{a'}}{\partial x^b} = \frac{\partial^2 x^{a'}}{\partial x^c\, \partial x^b} \frac{dx^c}{du} \overset{\text{(analytic)}}{=} \frac{\partial^2 x^{a'}}{\partial x^b\, \partial x^c} \frac{dx^c}{du} = X^{a'}_{bc} \frac{dx^c}{du},$$

we have that

$$\frac{d\lambda^{a'}}{du} = X^{a'}_{b} \frac{d\lambda^b}{du} + \lambda^b \frac{d}{du} X^{a'}_{b} = X^{a'}_{b} \frac{d\lambda^b}{du} + X^{a'}_{bc} \frac{dx^c}{du} \lambda^b. \tag{2.40}$$

The extra term prevents the object $\frac{d\lambda^a}{du}$ from being a vector. ■

One way to understand the problem is to recognize that the numerator of equation (2.41), $\lambda^a(u+\delta u) - \lambda^a(u)$, is the difference between a vector in the coordinate system at Q and a vector in the coordinate system at P. The difference is thus not a vector in either coordinate system.

Another way to understand the problem is that $X^{a'}_{\ b}$ depends on position:

$$(X^{a'}_{\ b})_u \neq (X^{a'}_{\ b})_{u+\delta u}\ .$$

Hence, $(X^{a'}_{\ b})_u\, \lambda^b\,(u+\delta u) - (X^{a'}_{\ b})_u\, \lambda^b\,(u) \neq (X^{a'}_{\ b})_{u+\delta u}\, \lambda^b\,(u+\delta u) - (X^{a'}_{\ b})_u\, \lambda^b\,(u)$. So, in the limit, there is the question of equality:

$$\frac{d\lambda^{a'}}{du} = \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\lambda^{a'}(u+\delta u) - \lambda^{a'}(u)}{\delta u} = \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\left(X^{a'}_{\ b}\right)_{u+\delta u}\lambda^b(u+\delta u) - \left(X^{a'}_{\ b}\right)_u \lambda^b(u)}{\delta u} \quad \text{and}$$

$$(X^{a'}_{\ b})_u\, \frac{d\lambda^b}{du} = \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\left(X^{a'}_{\ b}\right)_u \lambda^b(u+\delta u) - \left(X^{a'}_{\ b}\right)_u \lambda^b(u)}{\delta u}$$

Because of the extra term in equation (2.40), they turn out not to be equal.

Were both vectors in the difference defined at the same point, we could define a vector derivative that is a vector. Fortunately, $\lambda^a\,(u)$ generates two different vectors at Q: $\lambda^a\,(u+\delta u)$ and $\overline{\lambda}^a$, the parallel transport of $\lambda^a\,(u)$.

Definition The **absolute derivative of vector $\boldsymbol{\lambda^a\,(u)}$ along $\boldsymbol{\gamma}$** is

$$\boxed{\frac{D\lambda^a}{du} = \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\lambda^a(u+\delta u) - \overline{\lambda}^a}{\delta u}}\ .$$

It is a vector because it is a limit of vectors. As such, it transforms as a vector:

$$\frac{D\lambda^{a'}}{du} = \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\lambda^{a'}(u+\delta u) - \overline{\lambda}^{a'}}{\delta u} = \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\left(X^{a'}_{\ b}\right)_{u+\delta u}\lambda^b(u+\delta u) - \left(X^{a'}_{\ b}\right)_{u+\delta u}\overline{\lambda}^b}{\delta u}$$

$$= \left[\ \underset{\delta u\to 0}{\mathrm{Lim}} \left(X^{a'}_{\ b}\right)_{u+\delta u}\right] \left[\ \underset{\delta u\to 0}{\mathrm{Lim}}\ \frac{\lambda^b(u+\delta u) - \overline{\lambda}^b}{\delta u}\right] = \left(X^{a'}_{\ b}\right)_u \frac{D\lambda^b}{du} \qquad \checkmark$$

Next, we generate a formula for the absolute derivative. First, Taylor series are valid for any analytic function. So,

$$\lambda^a(u+\delta u) = \lambda^a(u) + \frac{d\lambda^a}{du}\,\delta u + \frac{1}{2!}\,\frac{d^2\lambda^a}{du^2}\,\delta u^2 + \ldots$$

and we see that to first order

$$\lambda^a(u+\delta u) \approx \lambda^a(u) + \frac{d\lambda^a}{du}\,\delta u. \qquad \text{(a)}$$

Hence,

$$\bar{\lambda}^a \overset{(2.27)}{\approx} \lambda^a(u) - \Gamma^a_{bc}(u)\, \lambda^b(u)\, \delta x^c$$

$$\frac{\lambda^a(u+\delta u) - \bar{\lambda}^a}{\delta u} \overset{(a)}{\approx} \frac{\frac{d\lambda^a}{du}\, \delta u + \Gamma^a_{bc}\, \lambda^b\, \delta x^c}{\delta u} = \frac{d\lambda^a}{du} + \Gamma^a_{bc}\, \lambda^b\, \frac{\delta x^c}{\delta u}$$

$$\boxed{\frac{D\lambda^a}{du} = \frac{d\lambda^a}{du} + \Gamma^a_{bc}\, \lambda^b\, \frac{dx^c}{du}} \tag{2.42}$$

$\lambda^a = \lambda^a(u)$ and $\Gamma^a_{bc} = \Gamma^a_{bc}(u)$ are evaluated at P. The absolute derivative is a vector even though neither term on RHS of equation (2.42) is. The absolute derivative is a combination of the total derivative and a term having connection coefficients (that connects P to nearby points in the difference quotient of the limit).

We observe that equation (2.23) for parallel transport of a contravariant vector along a curve can be written $\frac{D\lambda^a}{du} = 0$. We capture this fact as a theorem.

Theorem 2.3.2 $\lambda^a(u)$ form a parallel field of vectors along $\gamma$ if and only if $\frac{D\lambda^a}{du} = 0$.

We can extend the definition of absolute derivative to tensor fields $\tau^{a_1 \dots a_r}_{b_1 \dots b_s}(u)$ in two ways. We can extend the definition of parallelism between nearby tangent spaces $\mathbf{T}_P$ and $\mathbf{T}_Q$ to nearby type ($r$, $s$) tensor spaces at P and Q. This will lead to a formula for the absolute derivative $\frac{D}{du}$ as well as four conditions that $\frac{D}{du}$ must satisfy.

The other approach, the one we will use, is to define absolute derivative of a tensor to be the function that satisfies the four reasonable conditions plus equation (2.42). This will lead to a formula for absolute derivative of type ($r$, $s$) tensors, and parallel transport is then defined by requiring that the absolute derivative be zero. As with vectors, parallel transport is in general path dependent.

Definition The **absolute derivative of a tensor** is a function $\frac{D}{du}$ subject to the following conditions:

(a) $\frac{D}{du}$ applied to a type ($r$,$s$) tensor yields another type ($r$,$s$) tensor

(b) $\frac{D}{du}$ is a linear operation:

$$\frac{D\left(\kappa\tau^{a_1\ldots a_r}_{b_1\ldots b_s}\right)}{du} + \frac{D\left(\ell\sigma^{c_1\ldots c_r}_{d_1\ldots d_s}\right)}{du} + \;= k\,\frac{D\tau^{a_1\ldots a_r}_{b_1\ldots b_s}}{du} + \ell\,\frac{D\sigma^{c_1\ldots c_r}_{d_1\ldots d_s}}{du}$$

(c) $\frac{D}{du}$ obeys the Leibniz rule (aka product rule)

(d) For a scalar field, $\frac{D\varphi}{du} = \frac{d\varphi}{du}$

(e) For a vector field, D satisfies equation (2.42)

Definition Let $\gamma$ be a curve parameterized by $u$, and let $\tau = \tau^{a_1\ldots a_r}_{b_1\ldots b_s}$ be a tensor. We say that $\boxed{\boldsymbol{\tau(u)}\ \textbf{is generated by parallel transport along}\ \boldsymbol{\gamma}\ \text{if}\ \frac{D\tau}{du} = 0}$. This generalizes Definition (2.23) for parallel transport of vectors.

We develop equations for $\frac{D}{du}$ for lower rank tensors in order to uncover the pattern for general tensors. We list them here and then justify them. The “dot” notation represents differentiation with respect to the parameter $u$. For example, $\dot{\lambda}_a = \frac{d\lambda^a}{du}$.

| Field | Absolute Derivative Equation (along a curve) | |
|---|---|---|
| Scalar | $\boxed{\frac{D\varphi}{du} = \frac{d\varphi}{du}}$ | (2.44) |
| Contravariant vector | $\boxed{\frac{D\lambda^a}{du} = \dot{\lambda}^a + \Gamma^a_{c\,d}\,\lambda^c\,\dot{x}^d}$ | (2.45) |
| Covariant vector | $\boxed{\frac{D\mu_b}{du} = \dot{\mu}_b - \Gamma^c_{b\,d}\,\mu_c\,\dot{x}^d}$ | (2.46) |
| Type (2,0) tensor | $\boxed{\frac{D\tau^{a\,b}}{du} \equiv \dot{\tau}^{a\,b} + \Gamma^a_{c\,d}\,\tau^{c\,b}\,\dot{x}^d + \Gamma^b_{c\,d}\,\tau^{a\,c}\,\dot{x}^d}$ | (2.48) |
| Type (0,2) tensor | $\boxed{\frac{D\tau_{a\,b}}{du} \equiv \dot{\tau}_{a\,b} - \Gamma^c_{a\,d}\,\tau_{c\,b}\,\dot{x}^d - \Gamma^c_{b\,d}\,\tau_{a\,c}\,\dot{x}^d}$ | (2.49) |

Type (1,1) tensor $$\boxed{\frac{D\tau^a_b}{du} \equiv \dot{\tau}^a_b + \Gamma^a_{cd}\,\tau^c_b\,\dot{x}^d - \Gamma^c_{bd}\,\tau^a_c\,\dot{x}^d} \tag{2.50}$$

The formula (2.44) for scalar field is condition (d). The formula (2.45) for contravariant vector field is a repeat of equation (2.42). For the covariant vector formula (2.46), let $\mu_b$ be a covariant field along a curve $\gamma$, and let $\lambda^b$ be any contravariant field along $\gamma$. The total derivative may not yield a vector but it is defined in the usual way, unlike the absolute derivative. Hence, $\frac{d}{du}$ automatically obeys the product rule:

$$\frac{d\lambda^b}{du}\mu_b + \lambda^b\frac{d\mu_b}{du} \overset{\text{prod}}{\underset{\text{rule}}{=}} \frac{d(\lambda^b\mu_b)}{du} \overset{(2.44)}{=} \frac{D(\lambda^b\mu_b)}{du} \overset{(c)}{=} \frac{D\lambda^b}{du}\mu_b + \lambda^b\frac{D\mu_b}{du}$$

$$\overset{(2.45)}{=} \mu_b\left(\frac{d\lambda^b}{du} + \Gamma^b_{cd}\,\lambda^c\frac{dx^d}{du}\right) + \lambda^b\frac{D\mu_b}{du}$$

$$\Rightarrow \quad \lambda^b\frac{D\mu_b}{du} = \lambda^b\frac{d\mu_b}{du} - \Gamma^b_{cd}\,\lambda^c\frac{dx^d}{du}\mu_b \overset{c\leftrightarrow b}{=} \lambda^b\frac{d\mu_b}{du} - \Gamma^c_{bd}\,\lambda^b\frac{dx^d}{du}\mu_c$$

$$= \lambda^b\,(\dot{\mu}_b - \Gamma^c_{bd}\,\mu_c\dot{x}^d)$$

$\Leftrightarrow$ Equation (2.46) because it holds for arbitrary $\lambda^b$. ✓

As a guide for equation (2.48), we generate the formula for the special case where $\tau^{ab} = \lambda^a\mu^b$ and then make (2.48) the definition of the general case.

$$\frac{D\tau^{ab}}{du} = \frac{D\,\lambda^a\mu^b}{du} \overset{(c)}{=} \frac{D\,\lambda^a}{du}\mu^b + \lambda^a\frac{D\,\mu^b}{du}$$

$$\overset{(2.45)}{=} (\dot{\lambda}^a + \Gamma^a_{dc}\,\lambda^d\,\dot{x}^c)\,\mu^b + \lambda^a\,(\dot{\mu}^b + \Gamma^b_{dc}\,\mu^d\,\dot{x}^c)$$

$$= (\dot{\lambda}^a\,\mu^b + \lambda^a\,\dot{\mu}^b) + \Gamma^a_{dc}\,\lambda^d\,\mu^b\,\dot{x}^c + \Gamma^b_{dc}\,\lambda^a\,\mu^d\,\dot{x}^c$$

$$\overset{c\leftrightarrow d}{=} \frac{d\,\lambda^a\mu^b}{du} + \Gamma^a_{cd}\,\lambda^c\,\mu^b\,\dot{x}^d + \Gamma^b_{cd}\,\lambda^a\,\mu^c\,\dot{x}^d$$

$$= \dot{\tau}^{ab} + [\Gamma^a_{cd}\,\tau^{cb} + \Gamma^b_{cd}\,\tau^{ac}]\,\dot{x}^d \;✓$$

Equations (2.49) and (2.50) are Exercise 2.3.2. ✓

The mnemonic "co-below and minus" provides the pattern for equations (2.45 – 2.50) and leads to the general definition.

Definition The **absolute derivative of a type (*r*, *s*) tensor** is

$$\boxed{\frac{D\tau^{a_1\ldots a_r}_{b_1\ldots b_s}}{du} = \dot{\tau}^{a_1\ldots a_r}_{b_1\ldots b_s} + \left[\sum_{k=1}^{r} \Gamma^{a_k}_{c\,d}\,\tau^{a_1\ldots a_{k-1}\,c\,a_{k+1}\ldots a_r}_{b_1\ldots b_s} - \sum_{k=1}^{s} \Gamma^{c}_{b_k\,d}\,\tau^{a_1\ldots a_r}_{b_1\ldots b_{k-1}\,c\,b_{k+1}\ldots b_s}\right]\dot{x}^d} \tag{2.3-1}$$

For example, in $\tau^{a\,b}_{c}$ we use dummy variable $d$ in $\tau$ to replace, one-by-one, $a$, $b$, and $c$ and we use dummy variable $e$ with $\dot{x}$ :

Type (2,1) tensor: $\boxed{\dfrac{D\tau^{a}_{b\,c}}{du} = \dot{\tau}^{a\,b}_{c} + \left[\Gamma^{a}_{d\,e}\,\tau^{d\,b}_{c} + \Gamma^{b}_{d\,e}\,\tau^{a\,d}_{c} - \Gamma^{d}_{c\,e}\,\tau^{a\,b}_{d}\right]\dot{x}^e}$

## Covariant Derivative

We now extend the notion of derivative to a region U of a manifold M. We begin with a contravariant vector field $\lambda^a$ defined on U. If $\gamma$ is a curve in U, we first use equation (2.45) to define the absolute derivative of $\lambda^a$ restricted to $\gamma$:

$$\frac{D\lambda^a}{du} = \dot{\lambda}^a + \Gamma^a_{b\,c}\,\lambda^b\,\dot{x}^c\,. \tag{2.52}$$

Since $\dot{\lambda}^a = \dfrac{d\lambda^a}{du} = \dfrac{\partial\lambda^a}{\partial x^c}\dfrac{dx^c}{du} = \dfrac{\partial\lambda^a}{\partial x^c}\dot{x}^c$, we can rewrite this as

$$\frac{D\lambda^a}{du} = \left(\frac{\partial\lambda^a}{\partial x^c} + \Gamma^a_{b\,c}\,\lambda^b\right)\dot{x}^c\,. \tag{a}$$

The expression in the parentheses in equation (a) does not depend on $\gamma$ and, so, is a suitable candidate to be the derivative.

Claim $\tau^a_c = \dfrac{\partial\lambda^a}{\partial x^c} + \Gamma^a_{b\,c}\,\lambda^b$ is a (1,1) tensor:

In the language of the Quotient Theorem, $\tau^a_c$ is a "potential" (1,1) tensor. When contracted, we get

$$\tau^a_a = \frac{\partial\lambda^a}{\partial x^a} + \Gamma^a_{b\,a}\,\lambda^b$$

which is a scalar because all of the indices are contracted away. That is, $\tau^a_a$ is a sum of scalars. So, by the Quotient Theorem, $\tau^a_c$ is a type (1,1) tensor. ✓

We have shown that the expression in parentheses in equation (a) is a type (1,1) tensor, which justifies the following definition.

Definition The **covariant derivative of a vector field** $\lambda^a$ defined on a manifold region U is the type (1,1) tensor

$$\boxed{\lambda^a{}_{;c} \equiv \partial_c \lambda^a + \Gamma^a_{bc} \lambda^b = \lambda^a{}_{,c} + \Gamma^a_{bc} \lambda^b} \qquad (2.53)$$

where the **comma notation** is defined as $\boxed{\lambda^a{}_{,c} \equiv \dfrac{\partial \lambda^a}{\partial x^c}}$.

**Comma and semi-colon notation** are extended to repeated derivatives. For example:

$$\lambda^a{}_{,cb} = \partial_b \partial_c \lambda^a = \frac{\partial^2 \lambda^a}{\partial x^b \partial x^c}$$

$$\lambda^a{}_{;cb} = (\lambda^a{}_{;c})_{;b} = \partial_b (\partial_c \lambda^a + \Gamma^a_{dc} \lambda^d) + \Gamma^a_{eb} (\partial_c \lambda^e + \Gamma^e_{dc} \lambda^d)$$

Equation (2.53) is used as one of the conditions that define covariant derivatives of arbitrary tensors. After we specify conditions, we develop equations for lower rank tensors and unmask the general pattern for type ($r$,$s$) tensors.

Definition The **covariant derivative of a tensor** is a function, represented by a subscripted semi-colon, subject to conditions (A – E):

(A) The function applied to a type ($r$,$s$) tensor yields a type ($r$,$s$+1) tensor
(B) The function is a linear operator
(C) The function obeys the Leibniz rule (aka product rule)
(D) For a scalar field, $\varphi_{;a} = \partial_a \varphi$
(E) For a vector field, $\lambda^a{}_{;c} = \lambda^a{}_{,c} + \Gamma^a_{bc} \lambda^b$

| Field | Covariant Derivative Equation (for a manifold) | |
|---|---|---|
| Scalar | $\boxed{\varphi_{;a} = \partial_a \varphi}$ | (2.54) |
| Contravariant vector | $\boxed{\lambda^a{}_{;b} = \partial_b \lambda^a + \Gamma^a_{cb} \lambda^c}$ | (2.55) |
| Covariant vector | $\boxed{\mu_{a;c} = \partial_c \mu_a - \Gamma^b_{ac} \mu_b}$ | (2.56) |
| Type (2,0) tensor | $\boxed{\tau^{ab}{}_{;c} \equiv \partial_c \tau^{ab} + \Gamma^a_{dc} \tau^{db} + \Gamma^b_{dc} \tau^{ad}}$ | (2.57) |
| Type (0,2) tensor | $\boxed{\tau_{ab;c} \equiv \partial_c \tau_{ab} - \Gamma^d_{ac} \tau_{db} - \Gamma^d_{bc} \tau_{ad}}$ | (2.58) |

Type (1,1) tensor $\boxed{\tau^{a}_{b;c} \equiv \partial_c \tau^{a}_{b} + \Gamma^{a}_{dc}\, \tau^{d}_{b} - \Gamma^{d}_{bc}\, \tau^{a}_{d}}$ (2.59)

Type ($r$,$s$) tensor $\boxed{\begin{aligned}\tau^{a_1 \ldots a_r}_{b_1 \ldots b_s ;c} = \partial_c \tau^{a_1 \ldots a_r}_{b_1 \ldots b_s} &+ \textstyle\sum_{k=1}^{r} \Gamma^{a_k}_{dc}\, \tau^{a_1 \ldots a_{k-1}\, d\, a_{k+1} \ldots a_r}_{b_1 \ldots b_s} \\ &- \textstyle\sum_{k=1}^{s} \Gamma^{d}_{b_k c}\, \tau^{a_1 \ldots a_r}_{b_1 \ldots b_{k-1}\, d\, b_{k+1} \ldots b_s}\end{aligned}}$ (2.3-2)

Justification of these covariant derivative formulas closely mimics the justification given for the absolute derivative formulas. For example, to derive equation (2.56):

$$\mu_a\, \partial_b \lambda^a + \lambda^a\, \partial_b \mu_a = \partial_b(\lambda^a \mu_a) \overset{(2.54)}{=} (\lambda^a\, \mu_a)_{;b} \overset{(C)}{=} \mu_a\, \lambda^a{}_{;b} + \lambda^a\, \mu_{a;b}$$

$$\overset{(2.55)}{=} (\partial_b \lambda^a + \Gamma^a_{cb}\, \lambda^c)\, \mu_a + \lambda^a\, \mu_{a;b}$$

$$\Rightarrow \quad \lambda^a\, \mu_{a;b} = \lambda^a\, \partial_b \mu_a - \Gamma^a_{cb}\, \lambda^c\, \mu_a \overset{(a \leftrightarrow c)}{=} \lambda^a\, \partial_b \mu_a - \Gamma^c_{ab}\, \lambda^a\, \mu_c = \lambda^a\, (\partial_b \mu_a - \Gamma^c_{ab}\, \mu_c)\,.$$

Since this holds for all $\lambda^a$, and after exchanging $b \leftrightarrow c$, we get

$$\mu_{a;c} = \partial_c \mu_a - \Gamma^b_{ac}\, \mu_b \qquad ✓$$

Note 1 The mnemonic "co-below and minus" still applies.

Note 2 The partial and total derivatives of tensors do not obey the transformation laws and, so, do not generate tensors as do the absolute and covariant derivatives.

Note 3 Even if the order of partial differentiation doesn't matter, the order of covariant differentiation does. In other words, in general,

$$\lambda^a{}_{;bc} \neq \lambda^a{}_{;cb} \quad \text{even if} \quad \lambda^a{}_{,bc} = \lambda^a{}_{,cb}\,.$$

Theorem 2.3.3 The absolute and covariant derivatives of the metric tensor and Kronecker fields are zero. That is,

$$\boxed{g_{ab;c} = 0, \qquad \delta^a_{b;c} = 0, \qquad g^{ab}{}_{;c} = 0} \tag{2.60}$$

and along any curve $\gamma$ parameterized by $u$,

$$\boxed{\frac{\mathrm{D}g_{ab}}{\mathrm{d}u} = 0, \qquad \frac{\mathrm{D}\delta^a_b}{\mathrm{d}u} = 0, \qquad \frac{\mathrm{D}g^{ab}}{\mathrm{d}u} = 0}\,. \tag{2.61}$$

Proof.

$$0 \overset{(2.36)}{=} \partial_c g_{ab} - \Gamma_{bac} - \Gamma_{abc} \overset{(2.35)}{=} \partial_c g_{ab} - g_{bd}\,\Gamma^d_{ac} - g_{ad}\,\Gamma^d_{bc} \overset{(2.58)}{=} g_{ab;c} \quad ✓$$

$\partial_c \delta^a_b = 0$ because $\delta^a_b$ is a constant. So,

$$\delta^a_{b;c} \overset{(2.59)}{=} \partial_c \delta^a_b + \Gamma^a_{dc}\,\delta^d_b - \Gamma^d_{bc}\,\delta^a_d = 0 + \Gamma^a_{bc} - \Gamma^a_{bc} = 0 \quad ✓$$

$$0 \overset{(2.60)}{=} \delta^a_{b;c} = (g^{ad}\,g_{bd})_{;c} \overset{(C)}{=} g^{ad}{}_{;c}\,g_{db} + g^{ad}\,g_{db;c} \overset{(2.60)}{=} g^{ad}{}_{;c}\,g_{db} + 0 = g^{ad}{}_{;c}\,g_{db}.$$

$$0 = 0\,g^{be} = g^{ad}{}_{;c}\,g^{be}\,g_{db} = g^{ad}{}_{;c}\,\delta^e_d = g^{ae}{}_{;c} \;✓$$

Since $\dfrac{\partial}{\partial_u} = \dfrac{\partial}{\partial x^c}\dfrac{\partial x^c}{\partial_u} = \partial_u x^c\,\partial_c$,

$$\begin{aligned}\frac{Dg_{ab}}{du} &\overset{(2.49)}{=} \partial_u g_{ab} - \Gamma^c_{ad}\,g_{cb}\,\partial_u x^d - \Gamma^c_{bd}\,g_{ac}\,\partial_u x^d \\ &\overset{(c\leftrightarrow d)}{=} \partial_c g_{ab}\,\partial_u x^c - \Gamma^d_{ac}\,g_{db}\,\partial_u x^c - \Gamma^d_{bc}\,g_{ad}\,\partial_u x^c \\ &= (\partial_c\,g_{ab} - \Gamma^d_{ac}\,g_{db} - \Gamma^d_{bc}\,g_{ad})\,\partial_u x^c \\ &\overset{(2.58)}{=} g_{ab;c}\,\partial_u x^c \\ &\overset{(2.60)}{=} 0 \quad ✓\end{aligned}$$

The last two equalities are proven similarly. ∎

The next theorem states that inner products are preserved under parallel transport.

**Theorem 2.3.4** Suppose vector fields $\lambda^a$ and $\mu^a$ are parallelly transported along a curve $\gamma$. Then $\dfrac{d\left(g_{ab}\,\lambda^a\,\lambda^b\right)}{du} = 0$.

Proof. By Theorem 2.3.2, $\dfrac{D\lambda^a}{du} = \dfrac{D\mu^a}{du} = 0$. By equation (2.61), $\dfrac{Dg_{ab}}{du} = 0$. So,

$$\begin{aligned}\frac{d\left(g_{ab}\,\lambda^a\,\lambda^b\right)}{du} &\overset{(2.44)}{=} \frac{D\left(g_{ab}\,\lambda^a\,\lambda^b\right)}{du} \\ &\underset{\text{rule}}{\overset{\text{prod}}{=}} \lambda^a\,\mu^b\,\frac{Dg_{ab}}{du} + g_{ab}\,\mu^b\,\frac{D\lambda^a}{du} + g_{ab}\,\lambda^a\,\frac{D\mu^b}{du} = 0 \quad ∎\end{aligned}$$

Corollary 1 If vector fields $\lambda^a$ and $\mu^a$ are parallelly transported along $\gamma$ then their magnitudes as well as the angle between them remains constant.

Proof. By definition (1.80), their magnitudes remain constant. (We also know this from Theorem 2.2.2.) By definition (1.81), the angle between them remains constant. ■

Corollary 2 If $\lambda^a$ is parallelly transported along an affinely parameterized geodesic $\gamma$, it maintains a constant angle with the tangent vector at each point of the curve.

Proof. In Corollary 1, let $\mu^a$ be the tangent vector field. Since the tangent vector to an affinely parameterized geodesic (equation 2.12) satisfies the parallel transport equation (2.23), then the angle between $\lambda^a$ and $\mu^a$ remains constant by Corollary 1. ■

We finish this section with a discussion of divergence. Let a point $\boldsymbol{x}$ be enclosed in an infinitesimal surface S in Euclidean 3-space. Divergence at $\boldsymbol{x}$ is defined as the average magnitude of the components of $\boldsymbol{\lambda}$ outwardly normal to S. Recall that in the Cartesian coordinate system in Euclidean 3-space, **divergence** is defined mathematically as

$$\textbf{div}\ \boldsymbol{\lambda} \equiv \nabla \bullet \boldsymbol{\lambda} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \bullet (\lambda^x, \lambda^y, \lambda^z) = \frac{\partial \lambda^x}{\partial x} + \frac{\partial \lambda^y}{\partial y} + \frac{\partial \lambda^z}{\partial z} = \frac{\partial \lambda^i}{\partial x^i} = \lambda^i{}_{,i}\,. \quad (2.3\text{-}3)$$

If we assume an orthonormal natural basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ then by Theorem 1.2.6 (Exercise 1.2.4), $g_{ij} = \delta_{ij}$. Thus

$$\textbf{div}\ \boldsymbol{\lambda} \overset{(2.3-3)}{=} \lambda^i{}_{,i} = \delta_{ij} \frac{\partial \lambda^j}{\partial x^i} = \frac{\partial g_{ij} \lambda^j}{\partial x^i} = \frac{\partial \lambda_i}{\partial x^i} = \lambda_{i,i}\,.$$

Moreover, $\lambda^i{}_{;i}$ reduces to $\lambda^i{}_{,i}$:

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{i\ell} (\partial_j g_{k\ell} + \partial_k g_{\ell j} - \partial_\ell g_{jk}) = 0 \ \ \forall\, i, j, k$$

since $\partial_k g_{ij} = 0 \ \ \forall\, i, j, k$ in flat Euclidean space.

$$\Rightarrow \quad \lambda^i{}_{;i} \overset{(2.53)}{=} \lambda^i{}_{,i} \quad ✓$$

This is motivation for the following definitions.

## Definitions

**Divergence of a contravariant vector field** $\lambda^a$ is defined as the scalar field

$$\textbf{div } \lambda^a \equiv \nabla \bullet \boldsymbol{\lambda} = \lambda^a{}_{;a} . \tag{2.3-4}$$

**Divergence of a contravariant vector field** $\mu_a$ is defined as

$$\textbf{div } \mu_a \equiv \mu^a{}_{;a} \text{ where } \mu^a = \mu_b\, g^{a\,b} \text{ is the associated covariant vector.} \tag{2.3-5}$$

There are two **type (2,0) tensor field divergences**, defined by

$$\nabla \bullet \tau^{ab} \equiv \tau^{ab}{}_{;a} \text{ and } \textbf{div } \tau^{ab} \equiv \tau^{ab}{}_{;b} . \tag{2.3-6}$$

There are $(r + s)$ distinct **divergences for a type $(r,s)$ tensor field**, defined by

$$\textbf{div } \tau^{a_1 \dots a_r}_{b_1 \dots b_s} \equiv \tau^{a_1 \dots c \dots a_r}_{b_1 \dots b_s}{}_{;c} \text{ and } \textbf{div } \tau^{a_1 \dots a_r}_{b_1 \dots b_s} \equiv (\tau^{a_1 \dots a_r}_{b_1 \dots c \dots b_s}\, g^{c\,d})_{;d} . \tag{2.3-7}$$

Observations:

1. In the second of definitions (2.3-7), $g^{c\,d}$ is analogous to $g^{a\,b}$ in (2.3-5) that raises subscript *b,* but we do not specify how or if $g^{c\,d}$ raises subscript *c*.
2. If $\tau^{ab}$ is symmetric then the two definitions (2.3-6) coincide.

Example 2.3.1 Calculate divergence $\nabla \bullet \boldsymbol{r}$ in spherical coordinates.

From Example 1.1.4, where $u^1 = r,\ u^2 = \theta,\ u^3 = \phi$, we have that

$$\mathfrak{g} = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix} = r^4 \sin^2\theta$$

and the position vector is

$$r^i \boldsymbol{e}_i = \boldsymbol{r} \overset{(1.3)}{=} x\, \boldsymbol{i} + y\, \boldsymbol{j} + z\, \boldsymbol{k} = r \sin\theta \cos\phi\, \boldsymbol{i} + r \sin\theta \sin\phi\, \boldsymbol{j} + r \cos\phi\, \boldsymbol{k} = r\, \boldsymbol{e}_1$$

$$\Rightarrow \quad r^i = r\, \delta^i_1 .$$

$$\Rightarrow \quad \nabla \bullet \boldsymbol{r} \overset{(2.3-4)}{=} r^i{}_{;i} \overset{(2.55)}{=} \partial_i\, r^i + \Gamma^i_{j\,i}\, r^j = \partial_i\,(r\, \delta^i_1) + \Gamma^i_{i\,j}\, r\, \delta^j_1 = \frac{\partial r}{\partial r} + r\, \Gamma^i_{i\,1}$$

$$\overset{(2.37)}{=} 1 + \frac{1}{2}\, r\, \mathfrak{g}^{-1} \partial_1 \mathfrak{g} = 1 + \frac{1}{2}\, r\, \frac{1}{r^4 \sin^2\theta}\, 4\, r^3 \sin^2\theta = 1 + 2 = 3. \quad \blacksquare$$

## 2.4 Geodesic coordinates

In a coordinate system in which the metric tensor components are constant, the connection coefficients are zero in definition (2.13) for $\Gamma^a_{bc}$. This greatly simplifies

Parallel transport (2.23) : $\dfrac{d\lambda^a}{du} = 0$ for $a = 1 - N$ (2.4-1)

Absolute derivative (2.3-1): $\dfrac{D\tau^{a_1\ldots a_r}_{b_1\ldots b_s}}{du} = \dfrac{d\tau^{a_1\ldots a_r}_{b_1\ldots b_s}}{du}$ (2.4-2)

Covariant derivative (2.3-2): $\tau^{a_1\ldots a_r}_{b_1\ldots b_s\,;c} = \dfrac{\partial\tau^{a_1\ldots a_r}_{b_1\ldots b_s}}{\partial x^c}$ (2.4-3)

Euclidean space with Cartesian coordinates is such a system because $g_{ij} = \delta_{ij}$. While it is not possible to introduce such a system in a general curved manifold, it is possible at any given point, allowing simplified computations there.

Definition A coordinate system at a point O in a differentiable manifold M is known as a **geodesic coordinate system with origin O** if $\Gamma^a_{bc} = 0$ at O $\forall$ $a$, $b$, $c$.

Theorem 2.4.1 There is a geodesic coordinate system with origin O for any point.

Proof. Start with a coordinate system $x^a$ in which O has coordinates $x^a_O$ and connection coefficients $(\Gamma^a_{bc})_O$. Define a primed coordinate system

$$x^{a'} \equiv x^a - x^a_O + \frac{1}{2}(\Gamma^a_{bc})_O\,(x^b - x^b_O)\,(x^c - x^c_O). \qquad (2.62)$$

First, observe that point O is the origin in the primed coordinate system:

$$x^{a'}_O \overset{(2.62)}{=} x^a_O - x^a_O + \frac{1}{2}(\Gamma^a_{bc})_O\,(x^b_O - x^b_O)\,(x^c_O - x^c_O) = 0. \qquad ✓$$

Next,

$$X^{a'}_d \overset{(1.7-1)}{=} \frac{\partial x^{a'}}{\partial x^d} \overset{(2.62)}{=} \delta^a_d + \frac{1}{2}(\Gamma^a_{bc})_O\,[\,\delta^b_d\,(x^c - x^c_O) + \delta^c_d\,(x^b - x^b_O)\,]$$

$$= \delta^a_d + (\Gamma^a_{bc})_O\,\delta^b_d\,(x^c - x^c_O) = \delta^a_d + (\Gamma^a_{dc})_O\,(x^c - x^c_O) \qquad (a)$$

$$(X^{a'}_d)_O = \lim_{x^a\to x^a_O} X^{a'}_d \overset{(a)}{=} \delta^a_d, \quad \text{or} \quad \delta^d_c = (X^{d'}_c)_O \qquad (b)$$

$$\Rightarrow\ (\,(X^{a'}_{d})_O\ ) \overset{(b)}{=} \begin{pmatrix} \delta^1_1 & \delta^1_2 & \delta^1_3 \\ \delta^2_1 & \delta^2_2 & \delta^2_3 \\ \delta^3_1 & \delta^3_2 & \delta^3_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \det(\,(X^{a'}_{d})_O\,) = 1 \neq 0.$$

This means that $(\,(X^{a'}_{d})_O\,)$ is invertible, and so equation (2.62) satisfies the condition of an alternate system of coordinates in a neighborhood U′ of O in manifold M. So,

$(X^{d}_{a'})_O = \delta^d_a$ : (c)

$$(X^{d}_{a'})_O = (X^{c}_{a'})_O\, \delta^d_c \overset{(b)}{=} (X^{c}_{a'})_O\, (X^{d'}_{c})_O \overset{(1.69)}{=} \delta^d_a \qquad ✓$$

$X^{a'}_{ef} = (\Gamma^a_{fe})_O$ : (d)

$$X^{a'}_{ef} = \frac{\partial}{\partial x^e} X^{a'}_{f} \overset{(a)}{=} \frac{\partial}{\partial x^e}\left[\ \delta^a_f + (\Gamma^a_{fc})_O\,(x^c - x^c_O)\ \right] = 0 + (\Gamma^a_{fc})_O\, \partial_e x^c = (\Gamma^a_{fc})_O\, \delta^c_e$$
$$= (\Gamma^a_{fe})_O \qquad ✓$$

$\Rightarrow\ (X^{a'}_{ef})_O \overset{(d)}{=} (\Gamma^a_{ef})_O.$ (e)

Consequently, $(\Gamma^{a'}_{b'c'})_O = 0$ : (f)

$$(\Gamma^{a'}_{b'c'})_O \overset{(\text{Exercise } 2.2\times.5)}{=} (\Gamma^d_{ef})_O\, (X^{a'}_{d})_O\, (X^{e}_{b'})_O\, (X^{f}_{c'})_O - (X^{e}_{b'})_O\, (X^{f}_{c'})_O\, (X^{a'}_{ef})_O$$
$$\overset{(b,c,e)}{=} (\Gamma^d_{ef})_O\, \delta^a_d\, \delta^e_b\, \delta^f_c - \delta^e_b\, \delta^f_c\, (\Gamma^a_{ef})_O = (\Gamma^a_{bc})_O - (\Gamma^a_{bc})_O = 0 \qquad ■$$

Example (Not in book) Let the manifold M be the unit sphere in Euclidean 3-space and O be the point (1,0,0) on the *x*-axis. Develop the geodesic coordinate system for spherical coordinates and prove equation (f), that the primed connections are zero, using equation (2.32) from Exercise 2.2.4 instead of Exercise 2.2.5 as above.

The geodesic origin is $O = (r, \theta, \phi)_O = (1, \frac{\pi}{2}, 0)$, the unprimed spherical coordinates are $x^1 = r,\ x^2 = \theta,\ x^3 = \phi$, and

$$(\,g_{ab}\,) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}, \quad (\,g^{ab}\,) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1/r^2 \sin^2\theta \end{pmatrix}$$

$\Gamma^1_{22} \overset{(2.9)}{=} -r,\quad \Gamma^1_{33} = -r\sin\theta,\quad \Gamma^2_{12} = \Gamma^2_{21} = 1/r,\quad \Gamma^2_{33} = -\sin\theta\cos\theta,$

$\Gamma^3_{23} = \Gamma^3_{32} = \cot\theta,\quad \Gamma^a_{bc} = 0$ otherwise.

$(\Gamma^{1}_{22})_O = -1,\quad (\Gamma^{1}_{33})_O = -1,\quad (\Gamma^{2}_{12})_O = \left(\Gamma^{2}_{21}\right)_O = 1,\quad \left(\Gamma^{2}_{33}\right)_O = 0,$
$\left(\Gamma^{3}_{23}\right)_O = \left(\Gamma^{3}_{32}\right)_O = 0,\quad \Gamma^{a}_{b\,c} = 0$ otherwise.

To use Exercise 2.2.4, we need an expression for $(X^{a}_{e',\,d'})_O$. We use equation (2.62).

$$x^a \overset{(2.62)}{=} x^{a'} + x^a_O - \tfrac{1}{2}\,(\Gamma^a_{b\,c})_O\,(x^b - x^b_O)\,(x^c - x^c_O) \tag{g}$$

$$\Rightarrow\ X^a_{d'} \overset{(g)}{=} \delta^{a'}_{d'} - \tfrac{1}{2}\,(\Gamma^a_{b\,c})_O\,[\,X^b_{d'}(x^c - x^c_O) + X^c_{d'}\,(x^b - x^b_O)\,]$$

$$= \delta^{a'}_{d'} - (\Gamma^a_{b\,c})_O\,X^b_{d'}\,(x^c - x^c_O) \tag{h}$$

$$\Rightarrow\ X^a_{e',\,d'} \overset{(h)}{=} -\,(\Gamma^a_{b\,c})_O\,[\,X^b_{d'}\,X^c_{e'} + X^b_{e',\,d'}\,(x^c - x^c_O)\,] \tag{j}$$

$$\Rightarrow\ (X^a_{e',\,d'})_O \overset{(j)}{=} -\,(\Gamma^a_{b\,c})_O\,(X^b_{d'})_O\,(X^c_{e'})_O \overset{(c)}{=} -\,(\Gamma^a_{b\,c})_O\,\delta^b_d\,\delta^c_e = -\,(\Gamma^a_{d\,e})_O \tag{k}$$

We now compute one of the non-zero connection coefficients in the primed system to confirm that $(\Gamma^{a'}_{b',\,c'})_O = 0$:

$$\Gamma^{1'}_{2',\,2'} \overset{(2.32)}{=} \Gamma^d_{f\,g}\,X^{1'}_d\,X^f_{2'}\,X^g_{2'} + X^d_{2',\,2'}\,X^{1'}_d \tag{$\ell$}$$

$$(X^{1'}_d)_O \overset{(b)}{=} \delta^1_d,\quad (X^f_{2'})_O \overset{(c)}{=} \delta^f_2,\quad (X^g_{2'})_O \overset{(c)}{=} \delta^g_2,\quad (X^d_{2',\,2'})_O \overset{(k)}{=} -\,(\Gamma^d_{22})_O \tag{m}$$

$$(\Gamma^{1'}_{2',\,2'})_O \overset{(\ell)}{=} (\Gamma^d_{f\,g})_O\,(X^{1'}_d)_O\,(X^f_{2'})_O\,(X^g_{2'})_O + (X^d_{2',\,2'})_O\,(X^{1'}_d)_O$$

$$\overset{(m)}{=} (\Gamma^d_{f\,g})_O\,\delta^1_d\,\delta^f_2\,\delta^g_2 - (\Gamma^d_{22})_O\,\delta^1_d$$

$$= (\Gamma^1_{22})_O - (\Gamma^1_{22})_O = 0 \qquad \blacksquare$$

An immediate result of the next theorem is that about each point O of spacetime we can introduce a local coordinate system in which $x^a_O = 0$, $\Gamma^\mu_{\nu\sigma} \approx 0$, and

$$g_{\mu\nu} \approx \eta_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{2.65}$$

showing that general relativity locally looks like special relativity.

Theorem 2.4.2 At any point O on a differentiable manifold Σ there is a geodesic coordinate system at O where the metric tensor matrix $G_O = (g_{ab})_O$ is diagonal, having 1's and -1's on the diagonal.

Proof. Let $x^a$ be any coordinate system on Σ, and $x^{a'}$ the geodesic coordinate system at O defined by equation (2.62). Define a third coordinate system at O

$$x^{a''} = p^{a''}_{b'}\ x^{b'} \tag{2.63}$$

where $P = (p^{a''}_{b'}) = (p^a_b)$ is any nonsingular matrix. ($P$ must be nonsingular because it represents a transformation.) Claim O is the origin of the double-primed system:

$$x^{a''}_O \overset{(2.63)}{=} p^{a''}_{b'}\, x^{b'}_O = p^{a''}_{b'}\,(0) = 0\ . \quad ✓$$

Next, differentiation of equation (2.63) yields

$$X^{a''}_{c'} = p^{a''}_{b'}\ X^{b'}_{c'} = p^{a''}_{b'}\ \delta^{b'}_{c'} = p^{a''}_{c'} \tag{n}$$

and

$$X^{a''}_{a'\,c'} = 0\ . \tag{p}$$

Since $(\Gamma^{a'}_{b'\,c'})_O = 0$ for all $a$, $b$, and $c$,

$$(\Gamma^{a''}_{b''\,c''})_O \overset{\text{(Exercise 2.25)}}{=} (\Gamma^{d'}_{e'\,f'})_O\,(X^{a''}_{d'})_O\,(X^{e'}_{b''})_O\,(X^{f'}_{c''})_O - (X^{e'}_{b''})_O\,(X^{f'}_{c''})_O\,(X^{a''}_{e'\,f'})_O$$
$$\overset{(p)}{=} 0\ .$$

That is, $x^{a''}$ is a geodesic coordinate system with origin O. ✓

The metric tensor $g_{a'\,b'}$ transforms as

$$(g_{a''\,b''})_O = (g_{c'\,d'})_O \left(X^{c'}_{a''}\right)_O \left(X^{d'}_{b''}\right)_O \overset{(n)}{=} (g_{c'\,d'})_O \left(p^{c'}_{a''}\right)_O \left(p^{d'}_{b''}\right)_O\ .$$

The matrix version of this is $G''_O = P^{\mathrm{T}} G'_O\, P$. The transpose occurs because we must interchange the rows and columns of matrix $P$ when we multiply by it on the left.

It only remains to be shown that a nonsingular matrix $P$ can be chosen so that $P^{\mathrm{T}} G'_O\, P$ results in a diagonal (metric tensor) matrix whose entries are +1's and -1's. The book makes a vague reference that this is proven Birkhoff and Mac Lane, a 1977 modern algebra classic that I don't have. I develop my own proof of this, below.

The process involves three steps that I list here.

Step 1  Let $V$ be the tangent space at O. Any real symmetric matrix (such as $G$) has a collection of eigenvectors that form an orthonormal basis for $V$.

Step 2  Define $P$ to be a matrix whose columns are the eigenvectors of $G$ divided by the square root of their respective magnitudes. Then $P$ is an invertible matrix.

Step 3  $P^{\mathrm{T}} G'_O P$ is a diagonal matrix whose +1 and -1 diagonal entries are just the signs of the eigenvalues (i.e., +1 for positive eigenvalues, -1 for negative eigenvalues).

Besides being a major digression, one of the reasons the book may not have discussed this theorem is that the notation of matrix algebra, as used in Section 0, differs greatly from tensor notation. For example, a contravariant vector is denoted as a column vector $\boldsymbol{v}$ and a covariant vector as a row vector $\boldsymbol{w}^{\mathrm{T}}$ (transpose). The equivalent of the tensor inner product $g_{ab} v^a w^b = v^a w_a$ is $\boldsymbol{v}^{\mathrm{T}} \boldsymbol{w}$ and can be considered to be a row vector times a column vector. The inner product of column vectors is the dot product, $\boldsymbol{v}_a \bullet \boldsymbol{w}_b$. We stay with matrix algebra notation for this proof.

**Step 1**  Development of the orthonormal eigenvector basis is Theorem 0.2.

**Step 2**  Denote the eigenvectors as column vectors $\boldsymbol{v}_b = \begin{pmatrix} v_b^1 \\ \vdots \\ v_b^N \end{pmatrix}$. Then

$$\boldsymbol{v}_a^{\mathrm{T}} = \begin{pmatrix} v_a^1 & \cdots & v_a^N \end{pmatrix},$$

and so

$$\boldsymbol{v}_a^{\mathrm{T}} \boldsymbol{v}_b = v_a^1 v_b^1 + \cdots + v_a^N v_b^N = \boldsymbol{v}_a \bullet \boldsymbol{v}_b = \delta_{ab}$$

because $\boldsymbol{v}_a$ and $\boldsymbol{v}_b$ are orthogonal unit vectors.

Denote the eigenvalue of $\boldsymbol{v}_b$ as $\lambda_b$. Define $P = \begin{pmatrix} \cdots & \dfrac{\boldsymbol{v}_b}{\sqrt{|\lambda_b|}} & \cdots \end{pmatrix}$.

Claim: The inverse matrix of $P$ is $P^{-1} = \begin{pmatrix} \vdots \\ \sqrt{|\lambda_a|}\, \boldsymbol{v}_a^{\mathrm{T}} \\ \vdots \end{pmatrix}$:

$$P P^{-1} = \begin{pmatrix} & \vdots & \\ \cdots & \sqrt{|\lambda_a / \lambda_b|}\, \boldsymbol{v}_b \boldsymbol{v}_a^{\mathrm{T}} & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & \sqrt{|\lambda_a / \lambda_b|}\, \delta_{ab} & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \quad ✓$$

**Step 3** $G'_O\, \boldsymbol{v}_b = \lambda_b\, \boldsymbol{v}_b \;\Rightarrow\; G'_O \dfrac{\boldsymbol{v}_b}{\sqrt{|\lambda_b|}} = \dfrac{\lambda_b\, \boldsymbol{v}_b}{\sqrt{|\lambda_b|}} = \text{sign}(\lambda_b)\, \sqrt{|\lambda_b|}\; \boldsymbol{v}_b$ .

$$\Rightarrow \quad G'_O\, P = \left( \cdots\; G'_O \frac{\boldsymbol{v}_b}{\sqrt{|\lambda_b|}} \;\cdots \right) = \left( \cdots\; \text{sign}\,(\lambda_b)\, \sqrt{|\lambda_b|}\; \boldsymbol{v}_b \;\cdots \right).$$

$$P^{\mathrm{T}} = \begin{pmatrix} \vdots \\ \dfrac{\boldsymbol{v}_a^{\mathrm{T}}}{\sqrt{|\lambda_a|}} \\ \vdots \end{pmatrix}.$$

$$\Rightarrow \quad P^{\mathrm{T}} G'_O\, P = \begin{pmatrix} & \vdots & \\ \cdots & \dfrac{\boldsymbol{v}_a^{\mathrm{T}}}{\sqrt{|\lambda_a|}}\, \text{sign}(\lambda_b)\, \sqrt{|\lambda_b|}\; \boldsymbol{v}_b & \cdots \\ & \vdots & \end{pmatrix}$$

$$= \begin{pmatrix} & \vdots & \\ \cdots & \text{sign}(\lambda_b)\, \sqrt{|\lambda_b / \lambda_a|}\; \delta_{a\,b} & \cdots \\ & \vdots & \end{pmatrix} = \begin{pmatrix} \text{sign}\,(\lambda_1) & & \\ & \ddots & \\ & & \text{sign}\,(\lambda_N) \end{pmatrix} \qquad \blacksquare$$

While a positive definite metric tensor results in the metric tensor having all 1’s on the diagonal, an indefinite metric tensor will have a mix of +1’s and -1’s. The convention among physicists is to arrange for the +1’s to precede the -1’s. Most mathematicians use the opposite convention.

# Appendix C  Tensors and Manifolds

## C.2  Dual Spaces

The objective in this subsection is to define abstract dual spaces and show that the cotangent space is the dual space of the tangent space, and vice-versa.

Definition  Let **V** be an abstract $n$-dimensional vector space and

$\mathbf{V}^* = \{f : \mathbf{V} \to \mathbb{R} : f \text{ is linear}\}$

$\mathbf{V}^{**} = \{\omega : \mathbf{V}^* \to \mathbb{R} : \omega \text{ is linear}\}$

**V*** is called the **dual space of V**, and **V****, the dual space of **V***, is called the **2nd dual space of V**. While the members of **V** are referred to as **vectors**, the members of **V*** are called **dual vectors** or **covectors**. Members $\boldsymbol{v}$ of **V** are written in boldface, but members f of **V*** and $\omega$ of **V**** are functions and are not bolded.

Theorem C2.1  **V*** and **V**** are vector spaces under the natural definitions of the zero function, the additive inverse, addition, and scalar multiplication.

Proof.  Let $f$ and $g$ be functions in **V*** and $\alpha$ a scalar. For **V***, the natural definitions are:

Zero function: $0(\boldsymbol{v}) \equiv \mathbf{0}$ for all $\boldsymbol{v}$ (C2-1)

Additive inverse of $f$: $(-f)(\boldsymbol{v}) \equiv -f(\boldsymbol{v})$ (C2-2)

Addition: $(f + g)(\boldsymbol{v}) \equiv f(\boldsymbol{v}) + g(\boldsymbol{v})$ (C2-3)

Scalar Multiplication: $(\alpha f)(\boldsymbol{v}) \equiv \alpha\, f(\boldsymbol{v})$ (C2-4)

With these definitions, it is easy to show that all the conditions for **V*** to be a vector space are satisfied. Similarly for **V**** with:

Zero function: $0(f) \equiv \mathbf{0}$ for all $f$ (C2-5)

Additive inverse of $\omega$: $(-\omega)(f) \equiv -\omega(f)$ (C2-6)

Addition: $(\omega + v)(f) \equiv \omega(f) + v(f)$ (C2-7)

Scalar Multiplication: $(\alpha\omega)(f) \equiv \alpha\, \omega(f)$ ■ (C2-8)

Theorem C2.2  Let

$\mathcal{B} = \{\mathbf{e}_a : a = 1, \cdots, N\}$ be a basis for **V** and

$\mathcal{B}^* \equiv \{e^b : \mathbf{V} \to \mathbb{R} : e^b(\mathbf{e}_a) = \delta^b_a\}$. (C2-9)

Then $\mathcal{B}^*$ is a basis for **V***

Proof.  We need to show that $\{e^b\}$ is linearly independent and spans V*.

<u>Linearly independent</u>  $\lambda_b\, e^b = 0 \;\Rightarrow\; \lambda_a = \lambda_b \delta^b_a = \lambda_b\, e^b(\mathbf{e}_a) = 0$ for $a = 1, \cdots, n$ ✓

<u>Spans V*</u>  Let $f \in \mathbf{V}^*$.  $f(\mathbf{e}_a) \in \mathbb{R}$. Denote it as $f(\mathbf{e}_a) = \epsilon_a$ for $a = 1, \cdots, n$.

If $\boldsymbol{v} \in \mathbf{V}$, then $\boldsymbol{v} = \lambda^a \boldsymbol{e}_a$ since $\mathcal{B} = \{\boldsymbol{e}_a\}$ is a basis for $\mathbf{V}$. So, for all $\boldsymbol{v}$ we have

$$f(\mathbf{v}) = f(\lambda^a \boldsymbol{e}_a) \overset{(*)}{=} \lambda^a f(\boldsymbol{e}_a) = \lambda^a \epsilon_a = \lambda^a \epsilon_b \delta^b_a \overset{(C2-9)}{=} \lambda^a \epsilon_b e^b(\boldsymbol{e}_a) = \epsilon_b e^b(\lambda^a \boldsymbol{e}_a)$$
$$= \epsilon_b e^b(\mathbf{v})$$

$\Rightarrow$ $f = \epsilon_b e^b$ ; i.e., $f$ is a linear combination of the $e^b$. ✓ ■

The justification for step (*), above, is that, by definition of $\mathbf{V}^*$, $f$ is a linear function.

Example C2-1 Let $\mathbf{V}$ and $\mathbf{W}$ be $N$-dimensional vector spaces of column vectors and row vectors, respectively. Basis vectors of $\mathbf{V}$ and $\mathbf{W}$, respectively, are

$$\boldsymbol{e}_a = \begin{pmatrix} 0 \\ \vdots \\ 1_a \\ \vdots \\ 0 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{e}^b = (0 \ldots 1^b \ldots 0)\,.$$

Define a real-valued function $e^b$ on the basis vectors of $\mathbf{V}$ by

$$e^b(\boldsymbol{e}_a) = \delta^b_a \tag{C2-10}$$

By Theorem C2.2, $\mathcal{B}^* = \{e^b\}$ is a basis for $\mathbf{V}^*$ and, so, the 1-1 and onto mapping $e^b \mapsto \boldsymbol{e}^b$ generates an isomorphism from $\mathbf{V}^*$ to $\mathbf{W}$. This is the basis of the proof in the next theorem.

Theorem C2.3 $\mathbf{V}^{**}$ is (vector space) isomorphic to $\mathbf{V}$.

Proof For $\boldsymbol{v} \in \mathbf{V}$ define $\omega_{\mathbf{v}} \in \mathbf{V}^{**}$ by $\omega_{\mathbf{v}}: \mathbf{V}^* \to \mathbb{R}$: $\omega_{\mathbf{v}}(f) \equiv f(\mathbf{v})$.
Claim: $\omega_{\mathbf{v}}(f+g) = \omega_{\mathbf{v}}(f) + \omega_{\mathbf{v}}(g)$ and $\omega_{\mathbf{v}}(\alpha f) = \alpha\, \omega_{\mathbf{v}}(f)$:

$$\omega_{\mathbf{v}}(f+g) = (f+g)(\mathbf{v}) \overset{(C2-3)}{=} f(\mathbf{v}) + g(\mathbf{v}) = \omega_{\mathbf{v}}(f) + \omega_{\mathbf{v}}(g) \quad ✓$$
$$\omega_{\mathbf{v}}(\alpha f) = (\alpha f)(\mathbf{v}) \overset{(C2-4)}{=} \alpha\, f(\mathbf{v}) = \alpha\, \omega_{\mathbf{v}}(f) \quad ✓$$

Define $h$: $\boldsymbol{V} \to \mathbf{V}^{**}$: $h(\boldsymbol{v}) = \omega_{\mathbf{v}}$.
Claim: $h(\boldsymbol{v}+\boldsymbol{w}) = h(\boldsymbol{v}) + h(\boldsymbol{w})$ and $h(\alpha\boldsymbol{v}) = \alpha\, h(\boldsymbol{v})$:
Let $f \in \mathbf{V}^*$ and $\boldsymbol{v}$, $\boldsymbol{w} \in \boldsymbol{V}$. Define $\boldsymbol{x} = \boldsymbol{v}+\boldsymbol{w}$. Then

$$[h(\boldsymbol{v} + \boldsymbol{w})](f) = [h(\mathbf{x})](f) = \omega_{\mathbf{x}}(f) = f(\boldsymbol{x}) = f(\boldsymbol{v}+\boldsymbol{w}) \overset{(f \text{ is linear})}{=} f(\boldsymbol{v}) + f(\boldsymbol{w})$$
$$= \omega_{\mathbf{v}}(f) + \omega_{\mathbf{w}}(f) = h(\boldsymbol{v})(f) + h(\boldsymbol{w})(f) = [h(\boldsymbol{v}) + h(\boldsymbol{w})](f) \quad ✓$$
$$h(\alpha\boldsymbol{v})(f) = \omega_{\alpha\mathbf{v}}(f) = f(\alpha\boldsymbol{v}) \overset{(f \text{ is linear})}{=} \alpha\, f(\boldsymbol{v}) = \alpha\, \omega_{\mathbf{v}}(f) = \alpha\,[h(\boldsymbol{v})(f)] \quad ✓$$

So, $h$ is a linear map and, thus 1-1, from **V** onto **V****. That is, $h$ is an isomorphic mapping of **V** onto **V****. ■

Intuitively, this means that **V and V**** are interchangeable. We can consider a member of **V** to instead be a member of **V**** on-the-fly, and vice-versa. Vectors can be regarded as functions, and dual vectors (functions) can be regarded as vectors. We will use this fact in the tensor definition C3-9 where a covariant vector acts as a function of a contravariant vector and vice-versa.

Theorem C2.4 Row vectors can be regarded as the dual space of column vectors, and column vectors can be regarded as the dual space of row vectors

Proof. Let $\boldsymbol{V}^\top$ be the space of row vectors that are transposes of column vectors:

$$\boldsymbol{v}^\top = (v_1 \ldots v_N) \in \boldsymbol{V}^\top \quad \Leftrightarrow \quad \boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \in \mathbf{V}.$$

By Theorem C2.2, $\mathcal{B}^* = \{e^a\}$ is a basis for $\mathbf{V}^*$, the dual space for the column vectors **V**. By Example C2-1, $\boldsymbol{\mathcal{B}}^\top = \{\mathbf{e}^a\}$ is a basis for $\boldsymbol{V}^\top$, and so the mapping $e^a \mapsto \mathbf{e}^a$ defines an isomorphism between the dual space of **V** and the space $\boldsymbol{V}^\top$ of row vectors. The converse statement can be proven similarly. ■

In General Relativity, $\mathbf{T_P}^*$ is defined as the cotangent space, generated by the gradient vectors. In vector space theory, the same symbol $\mathbf{T_P}^*$ is defined as the dual of the tangent vector space. Fortunately, by Theorem C2.5 below, they are isomorphic, and we are free to change between them on-the-fly as needed.

Theorem C2.5 The cotangent space can be considered to be the dual of the tangent space, and, conversely, the tangent space can be considered to be the dual of the cotangent space.

Proof.

Tangent space: $\mathbf{T}_\mathrm{P} = \{\lambda^a \mathbf{e}_a\}$ (basis members are vectors, bolded)
Cotangent space: $\mathbf{T_P}^* = \{\lambda_a \mathbf{e}^a\}$ (basis members are vectors, bolded)
Dual to tangent space: $\mathbf{T_P}^* = \{\lambda_a e^a\}$ (basis members are functions, not bolded)

The mapping $\mathbf{e}^a \mapsto e^a$ of the basis vectors generates the isomorphism $\mathbf{T}_\mathrm{P}^* \cong \mathbf{T_P}^*$ ✓

Conversely, the tangent space $\mathbf{T}_\mathrm{P}$ can be regarded as a vector space of column vectors and the cotangent space $\mathbf{T}_\mathrm{P}{}^*$ can be regarded as a vector space of row vectors. By Theorem C2.4, the dual to a space of row vectors can be regarded a space of column vectors. Thus, $(\mathbf{T}_\mathrm{P}{}^*)^*$ can be regarded as space of column vectors. Since $\mathbf{T}_\mathrm{P}$ and $(\mathbf{T}_\mathrm{P}{}^*)^*$ are both isomorphic to the vector space of $N$-dimensional column vectors, they are isomorphic to each other. ■

Theorem C2.5 underscores that cotangent vectors can be regarded as functions of tangent vectors, and vice versa. This is key to the definition of tensor products.

Theorem C2.6 Let $\boldsymbol{\lambda} = \lambda^a \boldsymbol{e}_a$ be a tangent vector, $\boldsymbol{\mu} = \mu_b \boldsymbol{e}^b$ a cotangent vector, and $\lambda = \lambda^a e_a$ and $\mu = \mu_b e^b$ their functional equivalents. Then

$$\mu(\boldsymbol{\lambda}) = \lambda(\boldsymbol{\mu}) = \lambda^a \mu_a \tag{C2-11}$$

Proof. $e^b$ and $e_a$, respectively, were defined on basis vectors as $e^b(\boldsymbol{e}_a) = \delta_a^b$ and also $e_a(\boldsymbol{e}^b) = \delta_a^b$ . So,

$$\lambda^a \mu_a = \lambda^a \mu_b \delta_a^b = \lambda^a \mu_b \, e^b(\boldsymbol{e}_a) = \mu_b e^b(\lambda^a \boldsymbol{e}_a) = \mu(\boldsymbol{\lambda}) \quad ✓$$

$$\lambda^a \mu_b \delta_a^b = \lambda^a \mu_b \, e_a(\boldsymbol{e}^b) = \lambda^a e_a(\mu_b \boldsymbol{e}^b) = \lambda(\boldsymbol{\mu}) \quad ✓ \qquad ■$$

## C.3 Tensors

Definition The **product of two abstract vectors spaces** is
$$\mathbf{V} \times \mathbf{W} \equiv \{(\boldsymbol{v},\boldsymbol{w}): \boldsymbol{v} \in \mathbf{V},\, \boldsymbol{w} \in \mathbf{W}\}\ . \tag{C3-1}$$

The **tensor product of V and W** is
$$\mathbf{V} \otimes \mathbf{W} \equiv \{T: \mathbf{V}^* \times \mathbf{W}^* \to \mathbb{R}: T \text{ is bilinear }\} \tag{C3-2}$$

where **V*** is the dual vector space of linear functionals as defined in Appendix C.2.

**Bilinear** means that the function $T$ is linear in each component, separately:
$$T(\alpha\, \boldsymbol{v} + \beta\, \boldsymbol{u},\, \boldsymbol{w}) = \alpha\, T(\boldsymbol{v},\, \boldsymbol{w}) + \beta\, T(\boldsymbol{u},\, \boldsymbol{w}) \tag{C3-3}$$
$$T(\boldsymbol{v},\, \alpha\, \boldsymbol{w} + \beta\, \boldsymbol{x}) = \alpha\, T(\boldsymbol{v},\, \boldsymbol{w}) + \beta\, T(\boldsymbol{v},\, \boldsymbol{x})$$

In this definition of tensor product, **V** and **W** can be any vector spaces, and that includes the duals of vector spaces. Thus, the definition encompasses **V⊗W***, **V*⊗W**, and **V*⊗W***. For example, the tensor product definition for **V*** and **W*** is
$$\mathbf{V}^* \otimes \mathbf{W}^* \equiv \{T: \mathbf{V} \times \mathbf{W} \to \mathbb{R}: T \text{ is bilinear }\}. \tag{C3-4}$$

Theorem C3.1 The tensor product space $\mathbf{V} \otimes \mathbf{W}$ is a vector space.

Proof. From Theorem C2.1, we know that function spaces are vector spaces when using the natural definitions for the zero function, the additive inverse, addition, and scalar multiplication. It only remains to confirm that $0$, $-T$, $S+T$, and $\alpha T$ are also bilinear functions, and this is straight-forward. ■

In general relativity, definition (C3-2) is usually applied to tangent and cotangent spaces. For example, the tensor product of $\mathbf{T}_P$ and $\mathbf{T}_P^*$ is
$$\mathbf{T}_P \otimes \mathbf{T}_P^* = \{T: \mathbf{T}_P^* \times \mathbf{T}_P \to \mathbb{R}: T \text{ is bilinear }\}.$$

This definition of tensor product also encompasses $\mathbf{T}_P \otimes \mathbf{T}_P$, $\mathbf{T}_P^* \otimes \mathbf{T}_P$, $\mathbf{T}_P^* \otimes \mathbf{T}_P^*$, and even $\mathbf{T}_P \otimes \mathbf{T}_Q$ where Q is a different point of M.

In addition to defining tensor products of vector *spaces*, we also define tensor products of individual *vectors*. We keep in mind that when a vector is a function, it is not bolded. However, a vector as an argument of a function is a traditional vector and *is* bolded.

Definition  The **tensor product of a tangent basis vector with a cotangent basis vector** is the function

$$\boxed{e_a \otimes e^b : \mathbf{T}_\mathrm{P}^* \times \mathbf{T}_\mathrm{P} \to \mathbb{R} : e_a \otimes e^b(\mu_d\, \mathbf{e}^d,\ \lambda^c\, \mathbf{e}_c) \equiv \mu_a\, \lambda^b} \,. \qquad \text{(C3-5)}$$

Notice that a tangent vector operates on a cotangent vector in the first coordinate, and a cotangent vector operates on a tangent vector in the second coordinate. This shows that the tensor symbol $\otimes$ serves as a bookkeeper, preventing intermingling of the first and second coordinates. It essentially allows us to perform independent operations in two different vector spaces simultaneously. The only slight intermingling that can occur is that scalars can be brought out from one coordinate and then put back into the other coordinate: $T(\alpha \boldsymbol{v},\boldsymbol{w}) = \alpha T(\boldsymbol{v},\boldsymbol{w}) = T(\boldsymbol{v},\alpha \boldsymbol{w})$.

The basis tensor can be denoted more compactly as

$$\boxed{e_a^b \equiv e_a \otimes e^b}.$$

From equation (C2-10), we observe that

$$\boxed{e_a^b(\mathbf{e}^d,\ \mathbf{e}_c) = \delta_a^d\, \delta_c^b}. \qquad \text{(C3-6)}$$

It is a simple matter to show that $e_a \otimes e^b$ is bilinear.

Theorem C3.2  $\mathcal{B} \equiv \{e_a \otimes e^b\}$ is a basis for $\boldsymbol{T}_\mathrm{P} \otimes \boldsymbol{T}_\mathrm{P}^*$.

Proof.  Let $T \in \mathbf{T}_\mathrm{P} \otimes \mathbf{T}_\mathrm{P}^*$. We must show that $T$ is a linear combination of terms $e_a \otimes e^b$. Let $\boldsymbol{w}^* = \mu_a \mathbf{e}^a \in \mathbf{T}_\mathrm{P}^*$,  $\boldsymbol{v} = \lambda^b \mathbf{e}_b \in \mathbf{T}_\mathrm{P}$,  and  $\tau_b^a = T(\mathbf{e}^a,\ \mathbf{e}_b)$. Then  $T$: $\mathbf{T}_\mathrm{P}^* \times \mathbf{T}_\mathrm{P} \to \mathbb{R}$ and

$$T(\boldsymbol{w}^*,\boldsymbol{v}) = T(\mu_a \mathbf{e}^a,\lambda^b \mathbf{e}_b) = \mu_a\, \lambda^b\ T(\mathbf{e}^a,\ \mathbf{e}_b) = \mu_a\, \lambda^b\ \tau_b^a \overset{\text{(C3–5)}}{=} \tau_b^a\ \mathrm{e}_\mathrm{a} \otimes \mathrm{e}^\mathrm{b}(\mu_d\, \mathbf{e}^d,\ \lambda^c\, \mathbf{e}_c)$$
$$= \tau_b^a\ \mathrm{e}_\mathrm{a} \otimes \mathrm{e}^\mathrm{b}(\boldsymbol{w}^*,\boldsymbol{v}).$$

Since this holds for all vectors $\boldsymbol{w}^*$ and $\boldsymbol{v}$, then

$$T = \tau_b^a\ \mathrm{e}_\mathrm{a} \otimes \mathrm{e}^\mathrm{b} \qquad \blacksquare \qquad \text{(C3-7)}$$

Observe that $T$ is an outer product; it includes terms having $\mathrm{e}_\mathrm{a} \otimes \mathrm{e}^\mathrm{b}$ for every combination of $a$ and $b$.

Also, note that $T$ is a sum of terms $\tau^a_b\, e_a \otimes e^b$ and cannot in general be expressed as a singleton tensor product, like $k\, e_c \otimes e^d$. We conclude that even though $\mathcal{B}$ consists only of singleton tensor products ($e_a \otimes e^b$), it generates a vector space that is richer than just a space of singleton products.

Finally, where are the familiar tensors like $\tau^{a\,b}$, $\tau_{a\,b}$, and $\tau^a_{\;b\,c}$? Ignoring Einstein summation notation for a moment, the answer is that in general relativity, a singleton tensor $T = \lambda^a e_a \otimes \mu_b e^b = \lambda^a \mu_b\, e_a \otimes e^b$ is expressed as $T \overset{(C3-7)}{=} \tau^a_b \equiv \lambda^a \mu_b$ where the bases $e_a \otimes e^b$ are ignored. That is, the tensor $\tau^a_b$ that we are familiar with is actually a function, an element of a function vector space. Equation (C3-7) is its definition. The other familiar tensor expressions are also functions, generalizations of equation (C3-7), as developed below.

Definition Members of $\mathbf{T}_P \otimes \mathbf{T}_P{}^*$ that can be expressed as a single tensor product (rather than a sum of tensor products) are called **decomposable.** In quantum mechanics, members that can't be decomposed are called **entangled.**

Corollary $\dim(\mathbf{T}_P \otimes \mathbf{T}_P{}^*) = (\dim \mathbf{T}_P)\,(\dim \mathbf{T}_P{}^*)$ (C3-8)

The tensor product definition can be extended to include more than just two vector spaces. As before, this definition represents tensor products of vectors, covectors, and a mixture of vectors and covectors.

Definition Let $\mathbf{T} = \mathbf{T}_P$ be a tangent vector space and $\mathbf{T}^*$ its dual space. The **tensor product of *k* vector spaces T with $\ell$ covector spaces T*** is

$$\underbrace{\mathbf{T} \otimes \cdots \otimes \mathbf{T}}_{k\text{-times}} \otimes \underbrace{\mathbf{T}^* \otimes \cdots \otimes \mathbf{T}^*}_{\ell\text{-times}} = \{T\text{: } \underbrace{\mathbf{T}^* \times \cdots \times \mathbf{T}^*}_{k\text{-times}} \times \underbrace{\mathbf{T} \times \cdots \times \mathbf{T}}_{\ell\text{-times}} \to \mathbb{R}\text{: } T \text{ is multilinear}\} \quad \text{(C3-9)}$$

where **multilinear** is the generalization of bilinear that simply means that $T$ is linear in each vector and covector space separately.

Definition The members, $T$, of a tensor product (C3-9) are called **tensors**.

Definition **Basis tensors** are defined as the functions

$$\boxed{e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k} \equiv e_{a_1} \otimes \cdots \otimes e_{a_k} \otimes e^{b_1} \otimes \cdots \otimes e^{b_\ell}} \quad \text{(C3-10)}$$

It is straight-forward to generalize Theorem C3.2 to get that the collection of objects $e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k}$ is a basis and that a tensor (function) $T$ is a multilinear sum

$$\boxed{T = T^{a_1 \cdots a_k}_{b_1 \cdots b_\ell} \, e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k}} \tag{C3-11}$$

where

$$\boxed{e^{b_1 \cdots b_\ell}_{a_1 \cdots a_k}\left(\mathbf{e}^{d_1}, \cdots, \mathbf{e}^{d_k}, \mathbf{e}_{c_1}, \cdots, \mathbf{e}_{c_\ell}\right) = \delta^{d_1}_{a_1} \cdots \delta^{d_k}_{a_k} \, \delta^{b_1}_{c_1} \cdots \delta^{b_\ell}_{c_\ell}}. \tag{C3-12}$$

Tensors of the form $T^{a_1 \cdots a_k}$ are called **rank (*k*, 0) contravariant tensors**, $T_{b_1 \cdots b_\ell}$ are called **rank (0, $\ell$) covariant tensors**, and $T^{a_1 \cdots a_k}_{b_1 \cdots b_\ell}$ are called **rank (*k*, $\ell$) mixed tensors**. Rank (1,0) tensors are contravariant vectors and rank (0,1) tensors are covariant vectors. We also extend this definition to include **rank (0,0) tensors**, defined to be scalars. This is consistent because scalars are defined as functions on the coordinates of P. That is,

$\varphi : \mathrm{M} \to \mathbb{R} : \varphi(x^a) \equiv \varphi(x^1, \cdots, x^n) = k$, where $k \in \mathbb{R}$.

We note that the tensor product definition (C3-9), above could have been made for $k$ distinct vector spaces $\mathbf{T}^k$ and $\ell$ distinct (and unrelated) covector spaces $\mathbf{T}_\ell{}^*$, but the notation would have then gotten rather messy with upper and lower indices.

We have previously defined tensors as objects that transform according to the equation (1.73). Here we define tensors as *functions*, which are *objects*, but we make no mention of coordinate transformation. To be consistent, we must show that tensors, defined as functions, transform according to equation (1.73).

Theorem C3.3 Tensor coefficients, as defined above, transform according to

$$\tau^{a_1{}' \ldots a_r{}'}_{b_1{}' \ldots b_s{}'} = X^{a_1{}'}_{c_1} \cdots X^{a_r{}'}_{c_r} \, X^{d_1}_{b_1{}'} \cdots X^{d_s}_{b_s{}'} \, \tau^{c_1 \ldots c_r}_{d_1 \ldots d_s} \,. \tag{C3-13}$$

where

$$X^{a_m{}'}_{c_n} = \frac{\partial x^{a_m{}'}}{\partial x^{c_n}} \quad \text{and} \quad X^{d_n}_{b_m{}'} = \frac{\partial x^{d_n}}{\partial x^{b_m{}'}} \tag{C3-14}$$

Proof We provide a proof for $\tau^a_b$ that is easily extendable to $\tau^{c_1 \ldots c_r}_{d_1 \ldots d_s}$.

Let $(x^a)$ and $(x^{a'})$ be be overlapping coordinate systems at point P of manifold M.

Set

$T = \tau_b^a \, e_a^b = \tau_{b'}^{a'} \, e_{a'}^{b'} .$ (C3-15)

We have been restricting our attention to differential manifolds having metric tensors. So, as usual, we assume existence of a (functional) metric tensor $g_{ab} \equiv e_a \bullet e_b$.

We can express $e_a$ as a linear combination of basis functions $e_{b'}$:

$e_a = w_a^{b'} e_{b'}$ . (C3-16)

Claim $w_a^{b'} = \dfrac{\partial x^{b'}}{\partial x^a}$: (C3-17)

The key to proving this claim and, indeed, the theorem, is that the line element $\mathrm{d}s^2$ is required to be invariant under coordinate changes.

Define a vector $\mathrm{d}\boldsymbol{s}$ by $\mathrm{d}s^2 = \mathrm{d}\boldsymbol{s} \bullet \mathrm{d}\boldsymbol{s}$. Then

$\mathrm{d}s \bullet \mathrm{d}s = \mathrm{d}s^2 \overset{(1.83)}{=} g_{ab} \, \mathrm{d}x^a \, \mathrm{d}x^b = (e_a \bullet e_b) \, \mathrm{d}x^a \, \mathrm{d}x^b = (\mathrm{d}x^a \, e_a) \bullet (\mathrm{d}x^b \, e_b)$

$\Rightarrow \quad \mathrm{d}s = \mathrm{d}x^a e_a$

$\mathrm{d}s$ is also an invariant, so

$\mathrm{d}s = \mathrm{d}x^a \, e_a = \mathrm{d}x^{b'} e_{b'}$ (C3-18)

$\Rightarrow \quad \mathrm{d}x^a \, w_a^{b'} e_{b'} \overset{(C3-16)}{=} \mathrm{d}x^a \, e_a \overset{(C3-18)}{=} \mathrm{d}x^{b'} e_{b'} \overset{(\text{chain rule})}{=} \left( \dfrac{\partial x^{b'}}{\partial x^a} \mathrm{d}x^a \right) e_{b'}$

$\Rightarrow \quad w_a^{b'} = \dfrac{\partial x^{b'}}{\partial x^a}$ ✓

So,

$e_a \overset{(C3-16)}{=} w_a^{b'} e_{b'} \overset{(C3-17)}{=} \dfrac{\partial x^{b'}}{\partial x^a} \, e_{b'} \overset{(C3-14)}{=} X_a^{b'} e_{b'}$ . (C3-19)

Similarly we can define $\mathrm{d}x_a$ such that $\mathrm{d}s^2 = g^{ab} \, \mathrm{d}x_a \, \mathrm{d}x_b$ :

The vector $\mathrm{d}s$ can be expressed as a linear sum of the bases $e^a$, and we use this expression to define $\mathrm{d}x_a$:

$\mathrm{d}s \equiv \mathrm{d}x_a e^a$ .

Then the line element can be expressed as

$\mathrm{d}s^2 = \mathrm{d}s \bullet \mathrm{d}s = (\mathrm{d}x_a e^a) \bullet (\mathrm{d}x_b e^b) = (e^a \bullet e^b) \, \mathrm{d}x_a \, \mathrm{d}x_b = g^{ab} \, \mathrm{d}x_a \, \mathrm{d}x_b$ . ✓

Note that points ($x^a$) are always denoted with superscripts. However, derivation of the transformation equation for $e^a$ requires a subscripted d$x_a$. The definition above for d$x_a$ is that it is a coefficient, not a point. So, we are free to use subscripts (as we just did).

The derivation follows the derivation of equation (C3-19). We express the invariance of d$s$ as

$$\mathrm{d}s = \mathrm{d}x_a e^a = \mathrm{d}x_{b'} e^{b'} . \qquad \text{(C3-20)}$$

We write $e^b$ as a linear combination of basis vectors,

$$e^b = w^a_{b'}\, e^{b'} ,$$

find that

$$w^a_{b'} = X^a_{b'} ,$$

and then

$$e^b = X^b_{a'}\, e^{a'} . \qquad \text{(C3-21)}$$

Putting this altogether, we get the desired result:

$$\tau^{a'}_{b'}\, e^{b'}_{a'} \overset{\text{(C3-15)}}{=} \tau^a_b\, e^b_a \overset{\text{(C3-10)}}{=} \tau^a_b\, e_a \otimes e^b \overset{\text{(C3-19, C3-21)}}{=} \tau^a_b ( X^{b'}_a\, e_{b'}) \otimes (X^b_{a'} e^{a'})$$

$$= X^{b'}_a X^b_{a'}\, \tau^a_b\, e^{b'}_{a'}$$

$$\Rightarrow \quad \tau^{a'}_{b'} = X^{b'}_a X^b_{a'}\, \tau^a_b . \qquad \blacksquare$$

**Corollary** $\dfrac{\mathrm{d}x^{b'}}{\mathrm{d}x^a} = \dfrac{\partial x^{b'}}{\partial x^a}$

Proof We showed

$$\mathrm{d}x^a\, w^{b'}_a e_{b'} \overset{\text{(C3-16)}}{=} \mathrm{d}x^a\, e_a \overset{\text{(C3-18)}}{=} \mathrm{d}x^{b'} e_{b'}$$

$$\Rightarrow \quad \mathrm{d}x^a\, \frac{\partial x^{b'}}{\partial x^a} \overset{\text{(C3-17)}}{=} \mathrm{d}x^a\, w^{b'}_a = \mathrm{d}x^{b'}$$

Dividing both sides by d$x^a$ yields the result. $\blacksquare$

The LHS is *not* a derivative. It is the ratio of two coefficients. However, given that d$x^a$ is now a well-defined object as opposed to the nebulous infinitesimal distance of conventional calculus, it is reassuring that the fraction, which suggests differentiation, in fact equals the indicated partial derivative. We will provide more insight into d$x^a$ in Appendix D3.

# Appendix D  Coordinate-free Framework

## D.0 Introduction

We generally think of geometric objects as being points, lines, triangles, and squares and other polygons. However, the list also includes curves, cubes, spheres, planes, and higher dimensional objects. In addition, in modern GR, it also includes vectors, oriented parallelograms, and higher dimensional oriented and non-oriented objects.

In particular, modern GR views tensors as geometric objects having an existence independent of any coordinate system. This puts GR on a firmer footing that the prior view of GR be a coordinate-dependent theory.

Modern GR makes use of the theory of geometric algebra (a.k.a Clifford algebra) and the follow-on geometric calculus. Section D.1 gives a very brief, high-level introduction to geometric algebra where I borrowed heavily from John Denker's "Introduction to Clifford Algebra". Section D.2 give a likewise very brief introduction to geometric calculus where I borrowed heavily from David Bachman's  "A Geometric Approach to Differential Forms" .

## D.1 Geometric Algebra Overview

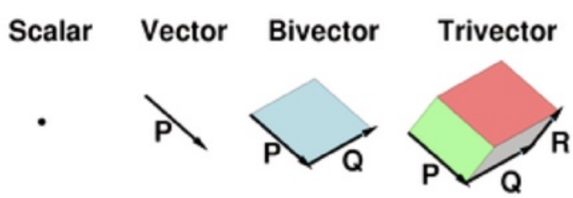



Figure D1-1  Blades

**Scalars** – In geometric algebra, "scalars" are points. A point can be viewed as a coordinate-free geometric object. If we draw two points on a piece of paper, we know which one is on the left and which is on the right, we can measure the distance, etc., all without a coordinate system. A scalar has no magnitude and represents zero dimensions. The scalars consist of all the real numbers. Geometrically, all numbers are viewed the same, as dimensionless points. They obey the familiar laws of arithmetic.

**Vectors** – These are coordinate-free objects that are visualized as having magnitude and direction. They are one dimensional.

**Bivectors** – The bivector generated by vectors *P* and *Q* can be visualized as the oriented parallelogram generated by drawing vectors head-to-tail (as shown above). The bivector *Q* followed by *P* would have the opposite orientation. Bivectors are two dimensional objects.

**Trivectors** – A trivector generated by three vectors is an oriented parallelopiped, as shown. The are six ways to combine the vectors *P*, *Q*, and *R*, but this results in only two orientations. *P*-*Q*-*R*, *Q*-*R*-*P*, and *R*-*P*-*Q* are all oriented the same way. *P*-*R*-*Q*, *Q*-*P*-*R*, and *R*-*Q*-*P* are oriented the opposite way.

⋮

**Blades** – Scalars, vectors, bivectors, trivectors, are all called **blades**. They are oriented hyper-parallelograms. One might expect the collection to be called multivectors, but that term can be confusing. For example, a trivector differs from a 3-vector, which is a single vector that resides in 3-space. John Denker uses the term multivector to mean a single vector in *k*-space, quite different than blades, which are hyper-parallelograms. Other books follow other naming conventions, including using the label multivector rather than blade.

**Grade of a blade** – The grade of a blade is the number of dimensions involved. Scalars are grade 0, vectors grade 1, bivectors grade 2, etc.

**Multiplication of blades by scalars** – If *V* is a vector, 2*V* is visualized as *V* with twice the length. Two times a bivector is visualized as a similarly-oriented parallelogram with twice the area. This paradigm works for all blades except scalars. Scalars have no size, so twice a scalar still has no size.

**Addition of blades** – Geometric algebra allows the addition of any two blades, regardless of whether they have the same grade. Adding blades that have different grades, say, a point and a bivector, may at first seem to be a strange idea. However, adding 1 + P is no stranger than adding 1 + 2*i*. A complex number *a* + *bi* is composed of two dissimilar objects, and so is the addition of blades that have different grades..

A general Clifford algebra element consists of the addition of blades: scalars, vectors, bivectors, trivectors, etc. In order to have a name for these objects, John Denker calls them **clifs**.  A clif is a sum of blades. The blades are subsets of the clifs just as the reals and imaginaries are subsets of the complex numbers.

Just as the space of complex numbers is far richer than the reals and imaginaries separately, the space clifs is much richer than the spaces of individual blades. Also, just as we never combine real and imaginary parts, we never combine clif parts of different grades.

As an example of a clif, let

$A = 1 + P,$

the sum of a scalar and a vector. Let

$B = 2 - 3P.$

When we add them we get the clif $A + B$:

$A + B = 3 - 2P.$

Given a clif, we can refer to the grade-0 piece, the grade-1 piece, the grade-2 piece, etc.

The first figure below shows how to add two vectors, head-to-tail, to acquire a new vector, something we already know how to do. The second figure expands the first figure to show that two bivectors can be added to obtain a new bivector, but only when certain sides match.

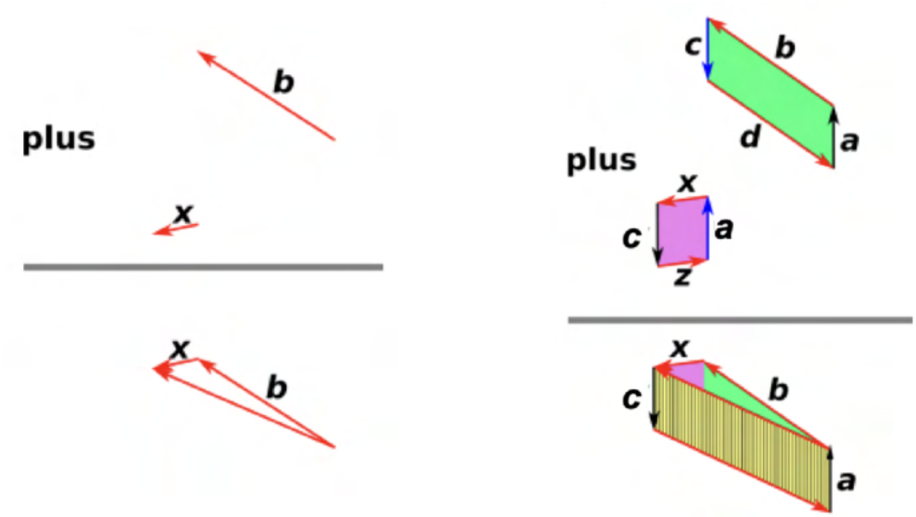



Figure D1-2 Blade addition

Had the sides not matched, we would have been left with a sum of disjoint bivectors, which is a grade 2 clif but not a blade. While the sum of any two vectors is a vector, this is not necessarily true for other blades. In general, a grade *k* clif is a sum of grade *k* blades.

In conventional calculus, there are two forms of vector multiplication: dot product and cross product. In geometric algebra, multiplication consists of geometric products, dot products, and wedge products. Wedge products replace cross product. Wedge products are easier to visualize and, unlike cross products, extend to any number of dimensions.

We conclude this short introduction by defining these products and showing that blades are wedge products.

**Product of blades** – We postulate a **geometric product** operator that can be used to multiply one clif by another clif. If *A* and *B* are clifs, we denote their geometric product by *AB*. We further postulate the geometric product is associative and distributes over addition:

$$(AB)C = A(BC) \equiv ABC$$
$$A(B+C) = AB + AC$$

M**ultiplication of scalars** – We further assume that scalars behave as usual:

- Product of scalars is just the usual commutative multiplication of real numbers
- The product of a scalar and a vector also represents the usual commutative multiplication
- In fact, we postulate that scalar multiplication is commutative with any blade

**Multiplication of vectors** – So far we have only postulated that vector multiplication exists, but not much else. However, for vectors *P* and *Q*, we *can* deduce that

$$PQ = \frac{PQ+QP}{2} + \frac{PQ-QP}{2}.$$

We use this to define **dot product** and **wedge product** of vectors:

$$P \bullet Q \equiv \frac{PQ+QP}{2} \tag{D1-1}$$

and

$$P \wedge Q \equiv \frac{PQ-QP}{2}. \tag{D1-2}$$

Wedge product is frequently called **exterior product**. Contrast this with the dot product that is widely called the **inner product**.

From equations (D1-1) and (D1-2), we immediately get the following properties for vectors:

$PQ = P{\bullet}Q + P{\wedge}Q$ (D1-3)

(geometric product is sum of dot product and wedge product)

$P \bullet Q = Q \bullet P$ (commutative or symmetric) (D1-4)

$P{\wedge}Q = -\, Q{\wedge}P$ (skew symmetric) (D1-5)

To justify the definition of dot product, we show that ***P•Q* is a scalar**:

Of the blades, only scalars are left unchanged when rotated by180°.
If we rotate a vector $P$ by 180°, we get $-P$, and so $(-P) \bullet (-P) = P{\bullet}P$. This shows that $P \bullet P$ is unchanged by a 180° rotation and is thus a scalar. Now let vector $R = P+Q$ and consider

$$R \bullet R = P \bullet P + 2P \bullet Q + Q \bullet Q.$$

Since $P$, $Q$, and $R$ are vectors, $P \bullet P$, $Q \bullet Q$, and $R \bullet R$ are scalars, and so the remaining term $P{\bullet}Q$ must be a scalar.
This shows that the dot product of any two vectors is a scalar. ✓

The formula for $P \bullet Q$ shows that for parallel vectors, $PQ = P \bullet Q$ and so $P{\wedge}Q = 0$.
We define two vectors to be orthogonal vectors if $P \bullet Q = 0$, and hence $PQ = P{\wedge}Q$.

Without providing the details, having formulas for dot and wedge products for both 90° and 180° rotations, we can perform various rotations both within and outside of the $PQ$-plane to show that $P{\wedge}Q$ behaves exactly as we would expect a bivector shown in Figure D1-1 to behave. We conclude that ***P∧Q* is a bivector**. Similarly, ***P∧Q∧R* is a trivector**, and so forth.

Also, without providing the details, a 4-dimensional manifold (such as spacetime) has blades of grades 0 to 4, but no higher. For example, the wedge product $\boldsymbol{A{\wedge}B{\wedge}C{\wedge}D{\wedge}E}$ of five vectors can be shown to be of grade one or three. So the clifford algebra for a 4-manifold consists of sums of scalars, vectors, bivectors, trivectors, and quadvectors.

Observe that the set of bivector blades is not a vector space because addition of two blades may not be another blade. So, the vector space generated by the blade includes sums of blades. Every vector space has a zero element, and even though 0 is a scalar, it is generated by the bivector blade $P{\wedge}P = 0$. So, the set of bivector blades does generate 0, and thus the blades do generate a vector space. In this vector space, the additive inverse of a bivector blade can be expressed as

$$-(P{\wedge}Q) \equiv (-P){\wedge}Q = P{\wedge}(-Q).$$

## D.2 Geometric Visualization of Tensors

Vectors, no matter how they are developed, are coordinate-free geometric objects. Since curves and surfaces are coordinate-free geometric objects, the following definition makes it clear that both contravariant and covariant vectors are coordinate-free geometric objects:

Definition A **contravariant vector** is defined as a tangent vector to a curve, and a **covariant vector** is defined as a gradient vector to a level surface.

As far as visualization goes, they are both just vectors, although when we have occasion to visualize them together, then they are perpendicular vectors.

We immediately get that the tangent space **T** and cotangent space **T*** are coordinate-free geometric objects.

We will collectively call tensors of ranks (1,0) and (0,1) to be rank 1 tensors, tensors of ranks (2,0), (1,1), and (0,2) to be rank 2 tensors, and so on.

Reviewing definition (C3-9), we can now observe that it is a coordinate-free definition of rank $(k,\ell)$ tensor $T$ as a multilinear functional:

$$T\colon \underbrace{\mathbf{T}^* \times \cdots \times \mathbf{T}^*}\times\underbrace{\mathbf{T} \times \cdots \times \mathbf{T}} \to \mathbb{R}.$$

Rank 1 tensors are vectors. Products of vectors are rank 2 tensors. In geometric algebra, we visualize these tensors as bivectors.

A sum $T$ of rank $k$ blade-tensors can be shown to be a rank $k$ tensor, a real-valued multilinear function. A sum of vector (blades) can be expressed as a vector (blade), so sums of vectors do not generate any new objects. But, a sum of bivector blades in general is a rank 2 tensor but not a bivector (blade).

In general, we make the following identifications of tensors with geometric objects:

| Tensor Rank | Blade | General Object | Tensor Rank | Blade | General Object |
|---|---|---|---|---|---|
| 0 | Scalar | Scalar | 3 | Trivector | Sum of trivectors |
| 1 | Vector | Vector | ⋮ | ⋮ | ⋮ |
| 2 | Bivector | Sum of bivectors | $N$ | $N$-blade | Sum of $N$-blades |

## D.3 Geometric Calculus

Geometric calculus expressions are much cleaner and easier-to-understand than their vector calculus counterparts. It is possible to solve equations which were previously impossible even to formulate due to their complexity.

Moreover, the vector cross product works only in 3-space, not in higher dimensions and not in lower dimensions. Equations involving cross products are a dead-end road. They cannot be extended to higher dimensions such as 4-dimensional spacetime.

The counterpart of cross product in geometric algebra is the bivector, which is fully extendable to both lower and higher dimensions. The result is that geometric calculus extends vector calculus to higher dimensions in ways not previously possible, enabling new equations and insights.

We begin with differential calculus and then move on to integral calculus. We conclude by translating some standard calculus theorems into geometric calculus notation. The comparisons will show the both the simplicity and clarity of geometric calculus.

Differentiation

The graph below shows the tangent line to $y = x^2$ at a point P in Euclidean coordinates. A vector at point P in has basis vectors $\boldsymbol{e_1}$ and $\boldsymbol{e_2}$, but no matter where P is located, $\boldsymbol{e_1}$ and $\boldsymbol{e_2}$ have the origin as base point. $\mathbb{R}^2$ is flat, and so it is convenient to view all vectors as emanating from the origin. However, the axes in a tangent plane $\mathbf{T}_P\mathbb{R}^2$ have base point P.

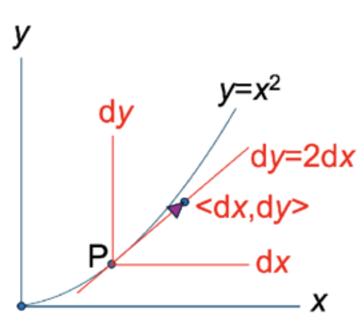



Figure D3-1 Differential axes

Geometric calculus uses d$x$ and d$y$ to label the axes in the tangent plane because the tangent space points do, after all, represent derivatives. While a line through the $\mathbb{R}^2$ base point (0,0) with slope 2 is written $y = 2x$ in the $xy$-plane, the line in the tangent plane through the $\mathbf{T}_P\mathbb{R}^2$ base point P=$\langle 0,0\rangle$ with slope 2 is written $\mathrm{d}y = 2\mathrm{d}x$. Points in $\mathbf{T}_P\mathbb{R}^2$ are labeled $\langle \mathrm{d}x,\mathrm{d}y\rangle$. Angle brackets are used to indicate that a point represents a vector. We will continue to use parentheses when representing coordinates and components.

The tangent space basis vectors are $\boldsymbol{e}_1 = \langle 1,0\rangle$ and $\boldsymbol{e}_2 = \langle 0,1\rangle$. Perhaps we should denote them $\mathrm{d}\boldsymbol{e}_1$ and $\mathrm{d}\boldsymbol{e}_2$ to distinguish them from the $\mathbb{E}^3$ basis vectors, but $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ is simpler notation and there should be no ambiguity as to which is being used.

Every point $\mathrm{d}\boldsymbol{x}$ in the tangent plane can be expressed as a linear combination of the basis vectors:

$$\mathrm{d}\boldsymbol{x} \equiv \langle \mathrm{d}x,\mathrm{d}y\rangle = \mathrm{d}x\,\langle 1,0\rangle + \mathrm{d}y\,\langle 0,1\rangle = \mathrm{d}x\,\boldsymbol{e}_1 + \mathrm{d}y\,\boldsymbol{e}_2\,. \qquad \text{(D3-1)}$$

With this approach, we see that d$x$ and d$y$ definitely do *not* represent the nebulous “infinitesimally small quantities” of conventional calculus. Rather, d$x$ and d$y$ simply represent the coordinates of points in the tangent plane.

There are several reasons why it makes sense to treat d$x$ and d$y$ as derivatives. The Corollary to Theorem C3.3 showed that the ratio $\frac{\mathrm{d}y}{\mathrm{d}x}$ can be treated as a derivative. Additionally, points Q in a tangent plane can be written parametrically, as

$$\mathrm{Q}(t) = \langle x_P^0 + t\,\mathrm{d}x,\ y_P^0 + t\,dy\rangle$$
$$\Rightarrow\quad \frac{\mathrm{dQ}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\langle x_P^0 + t\,\mathrm{d}x,\ y_P^0 + t\,\mathrm{d}y\rangle = \langle \mathrm{d}x,\ \mathrm{d}y\rangle.$$

This shows that the vector $\langle \mathrm{d}x,\mathrm{d}y\rangle$ is  a derivative.

In higher dimensions we can express equation (D3-1) as

$$\mathrm{d}\boldsymbol{x} = \mathrm{d}x^a\,\boldsymbol{e}_a\,, \qquad \text{(D3-2)}$$

and then we can express the vector $\mathrm{d}\boldsymbol{x}$ in terms of components:

$$\mathrm{d}\boldsymbol{x} = \langle \mathrm{d}x^a\rangle \qquad \text{(D3-3)}$$

## Integration

While the contravariant tangent plane vectors $\langle dx^a \rangle$ are suitable for differentiation, covariant tangent plane vectors, which are called 1-forms, are used for integration. We introduce 1-forms now.

Let $\mathbf{T}_P = \mathbf{T}_P M = \{\lambda^a \boldsymbol{e}_a\}$ be the tangent vector space at a point P of a Manifold M. The tangent space is bolded per the convention given in Appendix C.2.

Definition The coordinate-free definition of a 1-form is that a **1-form** is a linear function with a tangent vector as input and a real number as output. More precisely, a 1-form is a linear function from a tangent space at a point P to $\mathbb{R}$.

$$M: \quad \omega: \mathbf{T}_P M \to \mathbb{R} : \omega(\langle dx \rangle) = \omega_1 dx \tag{D3-4}$$

$$M^2: \quad \omega: \mathbf{T}_P M^2 \to \mathbb{R} : \omega(\langle dx, dy \rangle) = \omega_1 dx + \omega_2 dy \tag{D3-5}$$

$$\vdots \qquad \vdots$$

$$M^N: \quad \omega: \mathbf{T}_P M^N \to \mathbb{R} : \omega(\langle dx^a \rangle) = \omega_a dx^a \tag{D3-6}$$

Saying $\omega$ is a *linear* function means that

$$\omega(\alpha \langle dx^a \rangle + \beta \langle dy^a \rangle) = \alpha\, \omega(\langle dx^a \rangle) + \beta\, \omega(\langle dy^a \rangle). \tag{D3-7}$$

How do we get from the coordinate-free definition to the formulas (D3-4) to (D3-6)?

By definition, a function $\omega$ from $\mathbf{T}_P$ to $\mathbb{R}$ belongs to $\mathbf{T}_P^*$, the dual space of $\mathbf{T}_P$. The dual space has a basis $\{e^a\}$ defined as

$$e^a(\boldsymbol{e}_b) \overset{(C2-10)}{=} \delta_a^b .$$

Thus, in $M^N$ we can express $\omega$ as a linear combination of basis vectors,

$$\omega = \omega_a e^a,$$

which yields equation (D3-6):

$$\omega(\langle dx^a \rangle) = \omega_a e^a(dx^b \boldsymbol{e_b}) = \omega_a dx^b\, e^a(\boldsymbol{e}_b) = \omega_a dx^b \delta_b^a = \omega_a dx^a \quad \checkmark$$

As an alternative to equation (D3-6), we can regard the vector $\omega$ in terms of components:

$$\omega = \langle \omega_a \rangle. \tag{D3-8}$$

It can also be convenient to regard equation (D3-6) as a dot product of vectors:

$$\omega(d\boldsymbol{x}) = \omega \bullet d\boldsymbol{x}\text{:} \tag{D3-9}$$

$$\omega(d\boldsymbol{x}) \overset{(D3-6)}{=} \omega_a dx^a = \langle \omega_a \rangle \bullet \langle dx^a \rangle \overset{(D3-8,\, D3-3)}{=} \omega \bullet d\boldsymbol{x}$$

We can geometrically visualize a 1-form as its graph:

A 1-form on M is a line in $\mathbf{T}_P M \times \mathbb{R}$ through the origin: $z = \omega_1 dx$

A 1-form on $M^2$ is a plane through the origin in $\mathbf{T}_P M^2 \times \mathbb{R}$ : $z = \omega_1 dx + \omega_2 dy$

It is a surface because it has two parameters, $\omega_1$ and $\omega_2$.

It is a flat surface, a plane, because $z = \omega_1 dx + \omega_2 dy$ has no exponents.

It goes through the origin because $z = 0$ when $\omega_1$=0 and $\omega_2$=0.

⋮

A 1-form on $M^N$ is an $N$-dimensional hyperplane through the origin in $\mathbf{T}_P M^N \times \mathbb{R}$

We can also visualize a 1-form as a function that processes a tangent vector to obtain a real number. This process can be geometrically visualized. In $\mathbb{R}^1$, if the input to $\omega$ is the vector $\langle dx \rangle$, then the output is a vector, $\langle \langle dx \rangle, \omega_1 dx \rangle$, and the input vector is the downward projection of output vector. This is illustrated in the first graph below. The second graph represents $\mathbb{R}^2$. The input to $\omega$ is the vector $\langle dx, dy \rangle$, the output is the vector shown, and, again, the input vector is the downward projection of output vector.

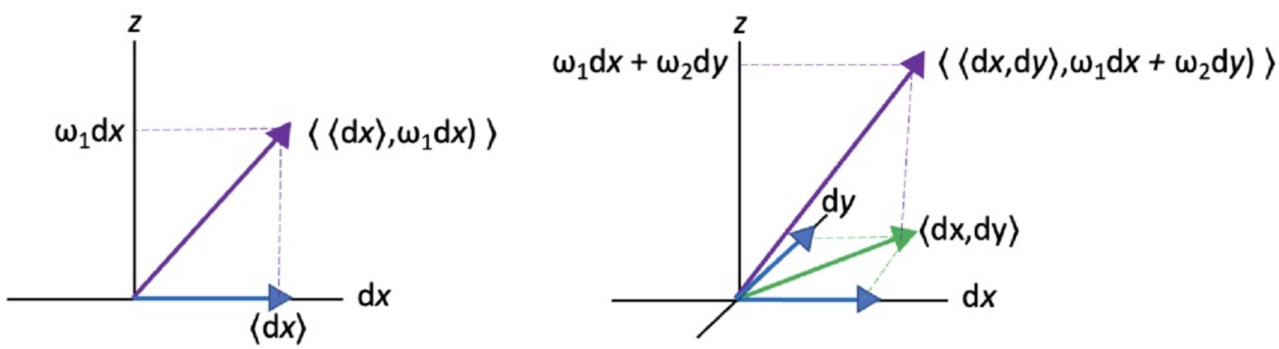



Figure D3.2 Output of one-forms

In the general case, both the input $\langle dx^a \rangle$ and the output $\langle \langle dx^a \rangle, \omega_a dx^a \rangle$ are visualized as vectors, with the latter projecting downward onto the former.

This can be geometrically interpreted in either of two ways. A 1-form is a function that generates a real number by:

1. Multiplying [Length of the projection of graph of 1-form onto the vector $\langle dx^a \rangle$] times [Magnitude of $\langle dx^a \rangle$] :

   This is the dot product interpretation of equation (D3-9):

   Multiply the length of the projection of one vector (the graph of $\omega$) onto another (the domain of $\omega$) by the magnitude of the second vector

2. Projecting the graph of $\omega$ onto each axis (to get $\langle dx^a \rangle$),
   then scaling by some constant ($\omega_a$),
   and then adding the terms ($\omega_a dx^a$) :
   This is simply putting equation (D3-6) into words

To introduce integration using 1-forms, we will imitate vector calculus in $\mathbb{R}^3$. Consider

| Vector | | 1-form | |
|---|---|---|---|
| $\boldsymbol{V} = F\boldsymbol{i} + G\boldsymbol{j} + H\boldsymbol{k}$ | $\leftrightarrow$ | $\omega = \langle F dx, G dy, H dz \rangle$ | (D3-10) |
| $\nabla = \frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j} + \frac{\partial}{\partial z}\boldsymbol{k}$ | $\leftrightarrow$ | $d = \langle \frac{\partial}{\partial x} dx, \frac{\partial}{\partial y} dy, \frac{\partial}{\partial z} dz \rangle$ | (D3-11) |
| $\nabla \bullet \boldsymbol{V} = \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z}$ | $\leftrightarrow$ | $d \bullet \omega = \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} + \frac{\partial H}{\partial z}$ | (D3-12) |
| $\nabla \times \boldsymbol{V}$ | $\leftrightarrow$ | $d\omega$ | |

To get the formulas for $\nabla \times \boldsymbol{V}$ and $d\omega$,

$$\nabla \times \boldsymbol{V} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F & G & H \end{vmatrix} = (H_y - G_z)\,\boldsymbol{i} + (F_z - H_x)\,\boldsymbol{j} + (G_x - F_y)\,\boldsymbol{k} \qquad \text{(D3-13)}$$

and

$$\begin{aligned} d\omega \equiv d \wedge \omega \overset{(D3-11,\, D3-10)}{=} & \frac{\partial F}{\partial x} dx \wedge dx + \frac{\partial F}{\partial y} dy \wedge dx + \frac{\partial F}{\partial z} dz \wedge dx \\ + \quad & \frac{\partial G}{\partial x} dx \wedge dy + \frac{\partial G}{\partial y} dy \wedge dy + \frac{\partial G}{\partial z} dz \wedge dy \\ + \quad & \frac{\partial H}{\partial x} dx \wedge dy + \frac{\partial H}{\partial y} dy \wedge dy + \frac{\partial H}{\partial z} dz \wedge dy \end{aligned}$$

$d\omega$ can be simplified. In Appendix D1, we showed that for vectors $P$ and $Q$:

$P \wedge Q$ is a bivector, visualized as a parallelogram generated by $P$ and $Q$,

$Q \wedge P \overset{(D1-5)}{=} - P \wedge Q$

$P \wedge P = 0$ because $P \wedge Q = 0$ for parallel vectors

$P \wedge Q = PQ$ if $P$ and $Q$ are perpendicular

Specifically,

$dx \wedge dx = dy \wedge dy = dz \wedge dz = 0$

$dy \wedge dx = - dx \wedge dy$, $dz \wedge dx = - dx \wedge dz$, and $dy \wedge dz = - dz \wedge dy$

So, $\mathrm{d}\omega$ matches the form of $\nabla \times \boldsymbol{V}$:

$$\mathrm{d}\omega = (H_y - G_z)\,\mathrm{d}y \wedge \mathrm{d}z \;+ (F_z - H_x)\,\mathrm{d}z \wedge \mathrm{d}x \;+ (G_x - F_y)\,\mathrm{d}x \wedge \mathrm{d}y\,.$$

Even better, since the vectors d$x$, d$y$, and d$z$ are perpendicular,

$$\mathrm{d}y \wedge \mathrm{d}z = \mathrm{d}y\,\mathrm{d}z, \quad \mathrm{d}z \wedge \mathrm{d}x = \mathrm{d}z\,\mathrm{d}x, \quad \text{and} \quad \mathrm{d}x \wedge \mathrm{d}y = \mathrm{d}x\,\mathrm{d}y\,,$$

and we get

$$\mathrm{d}\omega = (H_y - G_z)\,\mathrm{d}y\,\mathrm{d}z \;+ (F_z - H_x)\,\mathrm{d}z\,\mathrm{d}x \;+ (G_x - F_y)\,\mathrm{d}x\,\mathrm{d}y\,. \qquad \text{(D3-14)}$$

As an example of integration, we show that Green's theorem is nothing more than the Fundamental Theorem of Calculus. Recall Green's Theorem for P($x$,$y$) and Q($x$,$y$):

$$\int_C P\mathrm{d}x + Q\mathrm{d}y = \iint_D (Q_x - P_y)\,\mathrm{d}x\,\mathrm{d}y \qquad \text{(D3-15)}$$

where $D$ is a region in the $xy$-plane and the curve $C$ is its boundary. To use the $\mathbb{R}^3$ notation above, equations (D3-10 to D3-14), we set $\omega = P\mathrm{d}x + Q\mathrm{d}y$ and we identify $P$ with $F$ and $Q$ with $G$, and we impose the condition that changes in $z$ have no effect on $F$ and $G$. That is, $F_z = G_z = 0$. Thus $\mathrm{d}\omega = (Q_y - P_z)\,\mathrm{d}y\,\mathrm{d}z$, and Green's Theorem reads

$$\int_C \omega = \iint_D \mathrm{d}\omega.$$

In differential calculus we generally write $C$ as $\partial D$, the boundary of $D$, which gives Green's Theorem the formula

$$\int_{\partial D} \omega = \iint_D \mathrm{d}\omega. \qquad \text{(D3-16)}$$

To see that this is a generalization to 2 dimensions of the Fundamental Theorem of Calculus,

$$F(b) - F(a) = \int_{[a,b]} F'(x)\,\mathrm{d}x,$$

observe that the boundary of [$a$,$b$] consists of the points "a" and "$b$". Setting $\omega = \langle F\mathrm{d}x\rangle$ and, therefore, $\mathrm{d}\omega = \langle F'\mathrm{d}x\rangle$, leads to the geometric calculus version of the Fundamental Theorem of Calculus:

$$\int_{\partial[a,b]} \omega = F(b) - F(a)\,,$$

$$\int_{[a,b]} \mathrm{d}\omega = \int_{[a,b]} F'(x)\,\mathrm{d}x\,,$$

and

$$\int_{\partial[a,b]} \omega = \int_{[a,b]} \mathrm{d}\omega. \qquad \text{(D3-17)}$$

This is a 1-dimensional version of Green’s Theorem.

Note: Just as in Green’s theorem we use $\int_C P\mathrm{d}x + Q\mathrm{d}y$ as short-hand for $\int_C P(x,y)\,\mathrm{d}x + Q(x,y)\,\mathrm{d}y$, we use $\int_C \omega$ as short-hand for $\int_C \omega(\langle \mathrm{d}x, \mathrm{d}y \rangle)$.

We next introduce 2-forms, 3-forms, etc., and it turns out that several more of the important vector calculus formulas also become better understood as being nothing more than generalizations of the Fundamental Theorem of Calculus:

Stoke’s theorem where $S$ is a 3-dimensional surface and $C$ is its boundary:

$$\int_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{P} = \int_S (\nabla \times \boldsymbol{F}) \cdot \boldsymbol{n}\,\mathrm{d}\sigma$$

Since cross product cannot be generalized beyond 3 dimensions, this is end-of-the-line for Stoke’s Theorem in conventional calculus. In differential calculus, Stoke’s Theorem easily extends to any number of dimensions.

Stoke’s theorem where $\omega$ is a ($k$-1) form, d$\omega$ is a $k$-form, and M is a manifold:

$$\int_{\partial M} \omega = \iint_M \mathrm{d}\omega$$

Divergence theorem, flux of a vector field through a closed surface, where $R$ is a bounded set in $\mathbb{R}^3$ and the curve $S$ is its boundary:

$$\int_S (\nabla \times \boldsymbol{F}) \cdot \boldsymbol{n}\,\mathrm{d}\sigma = \int_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{P}$$

Divergence theorem where $\omega$ is a 2-form, d$\omega$ is a 3-form, and $R$ is a bounded set in $\mathbb{R}^3$:

$$\int_{\partial R} \omega = \iint_R \mathrm{d}\omega$$

Definition The coordinate-free definition of a 2-form is that a **2-form** is a bilinear function that has a pair of tangent vectors as input and a real number as output.

To convert this into a mathematical expression, we ask how a real number can be generated from a pair of tangent vectors? Recall from Appendix D1 that the wedge product of two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ is the bivector $\boldsymbol{v} \wedge \boldsymbol{w}$, which is visualized as the oriented parallelogram $\boldsymbol{vw}$. The area of a parallelogram is a real number and is what we choose to use for the definition the simplest 2-forms, ones that are products of 1-forms.

The next theorem gives a formula for the area of the bivector $\boldsymbol{vw}$ when $\boldsymbol{v}$ and $\boldsymbol{w}$ are vectors in $\mathbb{R}^2$. Because $\boldsymbol{vw}$ is an *oriented* parallelogram, we choose the output to be the *signed* area of the parallelogram, either positive or negative. Note that if $\boldsymbol{vw}$ is positively oriented, then $\boldsymbol{wv}$ is negatively oriented.

Theorem  Let $\boldsymbol{v} = \langle v_x, v_y \rangle$ and $\boldsymbol{w} = \langle w_x, w_y \rangle$. Then

Signed area of oriented parallelogram $\boldsymbol{vw}$

$$= \det \begin{pmatrix} v_x & w_x \\ v_y & w_y \end{pmatrix} = \begin{vmatrix} v_x & w_x \\ v_y & w_y \end{vmatrix} = v_x w_y - v_y w_x \, . \qquad \text{(D3-18)}$$

Proof  We first show theorem is true when $\boldsymbol{v}$ points along the $x$-axis

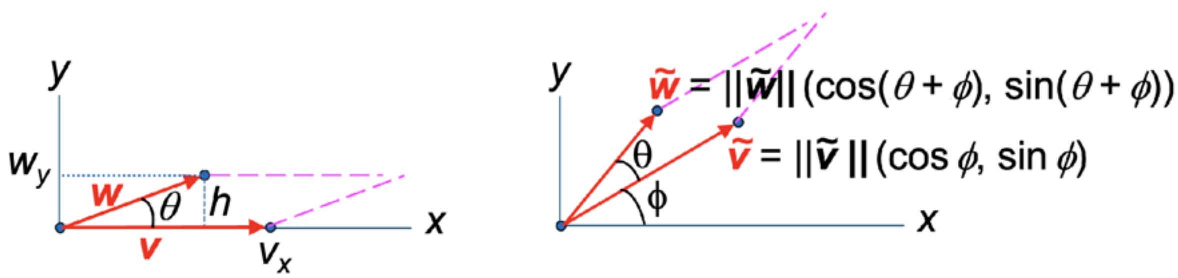



Figure D3.2  Signed area of parallelogram

Let $\theta$ be the angle between $\boldsymbol{vw}$ as shown in the figure on the left, where $0 \le \theta \le \pi$. So,

$$\sin\theta \ge 0. \qquad \text{(D3-19)}$$

From the left-hand figure, we see that the signed area of $\boldsymbol{vw}$ is

$$\text{Area} = (\text{base})\,(\text{height}) = v_x w_y = v_x w_y - v_y w_x \quad ✓ \qquad \text{(D3-20)}$$

since $v_y = w_x = 0$.

Also, observe from the left-hand figure that

$$v_x = \|\boldsymbol{v}\| \qquad \text{and} \qquad w_y = \|\boldsymbol{w}\| \sin\theta \qquad \text{(D3-21)}$$

The general case can be represented as a rotation of parallelogram $\boldsymbol{vw}$ by an angle $\phi$, as shown in the right-hand figure, where

$$\tilde{\boldsymbol{v}} = \|\boldsymbol{v}\|\,(\cos\phi, \sin\phi) \qquad \text{and} \qquad \tilde{\boldsymbol{w}} = \|\boldsymbol{w}\|\,(\cos(\theta+\phi), \sin(\theta+\phi)) \qquad \text{(D3-22)}$$

since $\|\tilde{\boldsymbol{v}}\| = \|\boldsymbol{v}\|$ and $\|\tilde{\boldsymbol{w}}\| = \|\boldsymbol{w}\|$.

We thus have

$\tilde{v}_x = \|\boldsymbol{v}\| \cos\phi$

$\tilde{v}_y = \|\boldsymbol{v}\| \sin\phi$

$\tilde{w}_x = \|\boldsymbol{w}\| (\cos\theta \cos\phi - \sin\theta \sin\phi)$

$\tilde{w}_y = \|\boldsymbol{w}\| (\sin\theta \cos\phi + \cos\theta \sin\phi)$ . (D3-23)

Therefore

$$\begin{vmatrix} \tilde{v}_x & \tilde{w}_x \\ \tilde{v}_y & \tilde{w}_y \end{vmatrix} = \tilde{v}_x \tilde{w}_y - \tilde{v}_y \tilde{w}_x$$

$$= \|\boldsymbol{v}\| \ \|\boldsymbol{w}\| \ | \sin\theta \cos^2\phi + \cos\theta \sin\phi \cos\phi - \cos\theta \sin\phi \cos\phi + \sin\theta \sin^2\phi \, |$$

$$= \|\boldsymbol{v}\| \ (\|\boldsymbol{w}\| \sin\theta) \overset{\text{(D3–21)}}{=} v_x \, w_y \overset{\text{(D3–20)}}{=} \text{signed area of } \boldsymbol{vw} = \text{signed area of } \tilde{\boldsymbol{v}}\tilde{\boldsymbol{w}} \qquad \blacksquare$$

We now have the formula for the output of the wedge product of 1-forms. If $\omega$ and $\nu$ are 1-forms on M:

$\omega, \nu : \mathbf{T}_P M \to \mathbb{R}$,

$\Rightarrow \quad \omega \wedge \nu : \mathbf{T}_P M \times \mathbf{T}_P M \to \mathbb{R}$ ,

$\Rightarrow \quad \langle\, \omega(\langle dx \rangle), \nu(\langle dx \rangle) \,\rangle$ and $\langle\, \omega(\langle du \rangle), \nu(\langle du \rangle) \,\rangle$ are vectors in $\mathbf{T}_P M \times \mathbf{T}_P M$.

Hence, the wedge product of these vectors is a bivector, a parallelogram:

$\langle\, \omega(\langle dx \rangle), \nu(\langle dx \rangle) \,\rangle \wedge \langle\, \omega(\langle du \rangle), \nu(\langle du \rangle) \,\rangle$. (D3-24)

Definition The **wedge product $\boldsymbol{\omega \wedge \nu}$** of 1-forms $\omega$ and $\nu$ is the function that has the signed area of parallelogram (D3-24) as output :

$$\omega \wedge \nu : \mathbf{T}_P M \times \mathbf{T}_P M \to \mathbb{R} : \omega \wedge \nu(\langle dx \rangle, \langle du \rangle) \equiv \begin{vmatrix} \omega(\langle dx \rangle) & \omega(\langle du \rangle) \\ \nu(\langle dx \rangle) & \nu(\langle du \rangle) \end{vmatrix} \qquad \text{(D3-25)}$$

Notes

1 It is straight-forward to show that $\omega \wedge \nu$ is a bilinear function and, hence, a 2-form.

2 Recall from Appendix D1 that we call a wedge product such as $\omega \wedge \nu$ a *blade*.

3 It is straight-forward to show that a sum of blade 2-forms is bilinear and, thus, also a 2-form.

In general, a sum of blade 2-forms cannot be represented as a single blade and, consequently, represent a sum of parallelograms but not a single parallelogram.

4 $\omega \wedge \nu$ is a rank (0,2) tensor, which is defined as a bilinear covariant function.

This definition of wedge product of 1-forms can be extended to product manifolds:

$\mathrm{M}^2$: $\omega \wedge \nu$: $\mathbf{T}_\mathrm{P}\mathrm{M}^2 \times \mathbf{T}_\mathrm{P}\mathrm{M}^2 \to \mathbb{R}$ : $\omega \wedge \nu(\langle \mathrm{d}x, \mathrm{d}u \rangle, \langle \mathrm{d}v, \mathrm{d}w \rangle)$

$$\equiv \begin{vmatrix} \omega(\langle \mathrm{d}x,\ \mathrm{d}u \rangle) & \omega(\langle \mathrm{d}v,\ \mathrm{d}w \rangle) \\ \nu(\langle \mathrm{d}x,\ \mathrm{d}u \rangle) & \nu(\langle \mathrm{d}v,\ \mathrm{d}w \rangle) \end{vmatrix} \qquad \text{(D3-26)}$$

⋮ ⋮

$$\mathrm{M}^N\text{:}\ \ \omega \wedge \nu\text{: } \mathbf{T}_\mathrm{P}\mathrm{M}^N \times \mathbf{T}_\mathrm{P}\mathrm{M}^N \to \mathbb{R} : \omega \wedge \nu(\langle \mathrm{d}x^a \rangle, \langle \mathrm{d}u^a \rangle) \equiv \begin{vmatrix} \omega(\langle \mathrm{d}x^a \rangle) & \omega(\langle \mathrm{d}u^a \rangle) \\ \nu(\langle \mathrm{d}x^a \rangle) & \nu(\langle \mathrm{d}u^a \rangle) \end{vmatrix} \qquad \text{(D3-27)}$$

Example Let $\omega = \langle 2\mathrm{d}x, -3\mathrm{d}y, \mathrm{d}z \rangle$ and $\nu = \langle \mathrm{d}x, 2\mathrm{d}y, -\mathrm{d}z \rangle$ be two 1-forms on $\mathbf{T}_\mathrm{P}\mathrm{M}^3$. Let $\mathrm{d}\boldsymbol{x} = \langle 1,3,1 \rangle$ and $\mathrm{d}\boldsymbol{u} = \langle 2,-1,3 \rangle$ be vectors in $\mathbf{T}_\mathrm{P}\mathrm{M}^3$. Then

$$\omega(\langle \mathrm{d}x^a \rangle) \overset{(D3-9)}{=} \langle \omega_a \rangle \bullet \langle \mathrm{d}x^a \rangle = 2 \cdot 1 - 3 \cdot 3 + 1 \cdot 1 = -6$$
$$\nu(\langle \mathrm{d}x^a \rangle) = \langle \nu_a \rangle \bullet \langle \mathrm{d}x^a \rangle = 1 \cdot 1 + 2 \cdot 3 - 1 \cdot 1 = 6$$
$$\omega(\langle \mathrm{d}u^a \rangle) = 10$$
$$\nu(\langle \mathrm{d}u^a \rangle) = -3$$
$$\omega \wedge \nu(\langle \mathrm{d}x^a \rangle, \langle \mathrm{d}u^a \rangle) = \begin{vmatrix} -6 & 10 \\ 6 & -3 \end{vmatrix} = -42,$$ the *signed* area of the parallelogram.

It is not difficult, using equation (D3-27), to show the following additional properties of the wedge operator:

1. $\omega \wedge \nu$ is skew symmetric: $\omega \wedge \nu(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{u}) = -\ \omega \wedge \nu(\mathrm{d}\boldsymbol{u}, \mathrm{d}\boldsymbol{x})$
2. $\wedge$ is skew symmetric: $\omega \wedge \nu = -\ \nu \wedge \omega$
3. $\omega \wedge \omega = 0$

We can geometrically visualize blade 2-forms as the graphs of functions:

In M, the graph is a line in $\mathbf{T}_\mathrm{P}\mathrm{M} \times \mathbf{T}_\mathrm{P}\mathrm{M} \times \mathbb{R}$ through the origin: $z = \omega \wedge \nu(\langle \mathrm{d}x \rangle, \langle \mathrm{d}u \rangle)$

In $M^2$, the graph is a plane through the origin in $\mathbf{T}_\mathrm{P}\mathrm{M}^2 \times \mathbf{T}_\mathrm{P}\mathrm{M}^2 \times \mathbb{R}$ :

$z = \omega \wedge \nu(\langle \mathrm{d}x, \mathrm{d}u \rangle, \langle \mathrm{d}v, \mathrm{d}w \rangle)$

It is a surface because it has two parameters, $\omega_1$ and $\omega_2$.

It is a flat surface, a plane, because $z = \omega_1 \mathrm{d}x + \omega_2 \mathrm{d}y$ has no exponents.

It goes through the origin because $z = 0$ when $\omega_1 = 0$ and $\omega_2 = 0$.

⋮

In $M^N$, the graph is a hyperplane through the origin in $\mathbf{T}_\mathrm{P}\mathrm{M}^N \times \mathbb{R}$

Like 1-forms, the output of 2-forms can also be visualized geometrically as projections of output vectors onto input vectors, but this is beyond the scope of what we wish to cover here.

We can also extend the blade 2-form pattern to define wedge product of multiple *1*-forms:

$$\omega_1\wedge\ldots\omega_N\colon \mathbf{T}_{\mathrm{P}}\mathrm{M}^N\times\ldots\mathbf{T}_{\mathrm{P}}\mathrm{M}^N \to \mathbb{R} : \omega_1\wedge\ldots\omega_N(\langle \mathrm{d}x^a\rangle,\ldots,\langle \mathrm{d}u^a\rangle)$$

$$\equiv \begin{vmatrix} \omega_1(\langle \mathrm{d}x^a\rangle) & \ldots & \omega_N(\langle \mathrm{d}u^a\rangle) \\ \vdots & & \vdots \\ \omega_N(\langle \mathrm{d}x^a\rangle) & \ldots & \omega_N(\langle \mathrm{du}^a\rangle) \end{vmatrix} \qquad \text{(D3-28)}$$

and it can be proven that determinant (D3-28) is the signed hypervolume of the resultant hyper-parallelopiped. $\omega_1\wedge\ldots\omega_N$ is an ***N*-form**. It is a blade, the simplest kind of *N*-form. Sums of *N*-form blades are also multilinear and, thus, *N*-forms.

In the remaining chapters, we develop GR following the Foster and Nightengale’s book, primarily using manifold basis vectors rather than tangent-plane basis vectors.

# Appendix A  Spacetime of Special Relativity

## A.0 Introduction

Notation  The book's equations are denoted (A.1), (A.2), ... My additional equations are numbered by section: (A0-1), (A0-2), ...., (A1-1), (A1-2), ...

This appendix develops special relativity using tensor mathematics. Coordinate-free development of GR has emerged during the decades since Foster and Nightengale published their book and was presented in Appendix C (not in the book). Having established this, we are free to continue with the book's development of GR using symbolic tensors as defined by the transformation equations.

The following definition of inertial frame, following Scott Hughes online MIT lecture on GR and also Blandford and Thorne, "Relativity and Cosmology, June 2021, is an elaboration of Foster and Nightengale's definition.

Definition  An **inertial frame** is a (conceptual) lattice of clocks and measuring rods that allows assignment of a coordinate system ($t$, $x$, $y$, $z$) with mutually orthogonal space coordinates and for which a particle in motion remains in motion unless disturbed by an outside force (i.e., Newton's first law holds). Specifically, the lattice properties are:

1. The lattice moves freely thought spacetime – no forces act on it and it does not rotate relative to distant beacons
2. The measuring rods are orthogonal and uniformly ticked, forming an orthonormal, Cartesian coordinate system
3. All of the clocks tick uniformly
4. The clocks are synchronized using the "Einstein synchronization procedure": clock 1 emits a pulse of light at $t = t_e$ that bounces off of a mirror at clock 2 and is received back at $t = t_r$. Both clocks are synchronized to report the same time of bounce: $t_b = \frac{t_e + t_r}{2}$.

If K and K′ are inertial frames, each frame may have velocity with respect to the other frame. Also, the space coordinates may be rotated with respect to each other, but it must be a fixed rotation, unchanging over time.

Definition  The **origin** of an inertial frame has coordinates (0,0,0,0).The **center** of an inertial frame is the space point (0,0,0).

An inertial frame has just a single origin, but the coordinates of its center as viewed by an observer in different inertial frame can vary over time .

Since the space coordinates are mutually orthogonal, inertial frames can be described using Euclidean coordinates although spherical and other coordinate systems are also permissible. We restrict treatment to Euclidean coordinate systems in this appendix. We note that spherical coordinates will play a large role when we develop the Schwarzschild solution to Einstein's GR field equations.

Because K and K′ may have different *xyz*- and x′y′z′-axes, $\boldsymbol{i}'$, $\boldsymbol{j}'$, and $\boldsymbol{k}'$ may not be the same as $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$.

## Postulates of Spacetime Relativity

1. The speed of light is the same in all inertial frames:

$$c = \frac{dr}{dt} = \frac{dr'}{dt'} \qquad \text{(A.1)}$$

   where $\boldsymbol{r} = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$ is the Euclidean position vector for space,
   $r^2 = |\,\boldsymbol{r}\,|^2 = \boldsymbol{r} \bullet \boldsymbol{r} = x^2 + y^2 + z^2$,
   $\boldsymbol{r}' = x'\boldsymbol{i}' + y'\boldsymbol{j}' + z'\boldsymbol{k}'$,
   and $r'^2 = x'^2 + y'^2 + z'^2$.

2. The laws of nature are the same in all inertial frames.

We will develop Lorentz transformations as the mathematical expression of how tensors transform between inertial frames. The "laws of nature" in Postulate 2 refer to physical properties that can be represented by tensors whose form is unchanged under Lorentz transformations.

As for Postulate 1, we can express equation (A.1) for photon speed as

$$c^2\,dt^2 - dr^2 = 0 = c^2\,dt'^2 - dr'^2\,. \qquad \text{(A0-1)}$$

where $dr^2 = dx^2 + dy^2 + dz^2$ and $dr'^2 = dx'^2 + dy'^2 + dz'^2$ .

Even though Postulate 1 deals with *photon* motion, it has implications for *particle* motion since speed of light, $c$, is the limit of particle speed. In particular, note that Equation (A0-1) causes time (d$t$) and space (d$r$) to have opposite signs, as implemented in the following definition for *particles*.

Definitions The path of a particle in spacetime is called its **world line**. A point $\mathcal{P} = (t, x, y, z)$ on the world line is called an **event.** A **displacement vector** between events $\mathcal{P}$ and $\mathcal{Q}$ is defined as

$$\Delta\vec{s} = (t_{\mathbb{Q}} - t_{\mathbb{P}}, x_{\mathbb{Q}} - x_{\mathbb{P}}, y_{\mathbb{Q}} - y_{\mathbb{P}}, z_{\mathbb{Q}} - z_{\mathbb{P}}) = (c\Delta t, \Delta x, \Delta y, \Delta z) = (c\Delta t, \Delta\mathbf{r}).$$

The **interval between two events is denoted $\Delta s$** and is defined as

$$\pm\,\Delta s^2 \equiv \Delta\,\vec{s}\bullet\Delta\,\vec{s} = c^2\,\Delta t^2 - \Delta r^2.$$

In the limit as $t$ goes to zero, $\Delta s$ goes to **d$s$, defined by**

$$\pm\,\mathrm{d}s^2 \equiv c^2\,\mathrm{d}t^2 - \mathrm{d}r^2 = c^2\,\mathrm{d}t'^2 - \mathrm{d}r'^2\;. \tag{A.2}$$

This shows that d$s$ is frame invariant. We call ***ds*** an **invariant interval between neighboring events**. Physicists generally use the plus sign while mathematician usually use the minus sign. We will follow the physics convention and set

$$\mathrm{d}s^2 \equiv\; c^2\,\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \;= c^2\,\mathrm{d}t'^2 - \mathrm{d}x'^2 - \mathrm{d}y'^2 - \mathrm{d}z'^2\;. \tag{A0-2}$$

Equation (A0-2) shows that, as with photons, the particle invariant interval is the same in all inertial frames. However, while d$s$ = 0 for light (because of Postulate 1, shown above), d$s$ is generally not zero for particles.

While each inertial frame has its own time, there is also an invariant "proper time" that represents a clock moving with the particle along its world line. The **proper time interval** between events is denoted **$\Delta\tau$** and defined as

$$c\,\Delta\tau \;\equiv\; \Delta s.$$

In the limit, the proper time interval is **d$\tau$**, expressed by

$$c\,\mathrm{d}\tau \;\equiv\; \mathrm{d}s. \tag{A0-3}$$

Foster and Nightingale are careful never to refer to $\frac{\mathrm{d}s}{\mathrm{d}\tau}$, perhaps because it indicates that the **"**proper speed" of a particle along its world line is $\frac{\mathrm{d}s}{\mathrm{d}\tau} = c$. Instead, a more useful velocity, called "world velocity", is defined in Section A.6 in terms of a particle's coordinates.

Notation $\boxed{x^0 \equiv ct,\quad x^1 \equiv x,\quad x^2 \equiv y,\quad x^3 \equiv z}$, or $\boxed{x^\mu = (c\,t,\,x,\,y,\,z)}$ (A.3)

Equation (A.3) is also expressed as $x^\mu = (x^0, x^1, x^2, x^3) = (ct, \boldsymbol{x})$ where $\boldsymbol{x} = (x, y, z)$. Using Euclidean coordinates, we also write $\boldsymbol{x}$ as a position vector $\boldsymbol{x} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$.

In Special and General Relativity, we use $\mu, \nu, \sigma, \rho = 0 - 3$ rather than $a, b, c$ (manifolds), and $i, j, k = 1 - 3$ for the spatial coordinates. $x^0$ is defined as $ct$, not just t, so that all the coordinate axes represent distance.

This notation enables us to rewrite equation (A.2) yet again, in a way that emphasizes that the invariant interval represents a unit of distance:

$$\boxed{ds^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu = \eta_{\mu'\nu'}\, dx^{\mu'}\, dx^{\nu'}} = c^2\, dt^2 - dx^2 - dy^2 - dz^2 \tag{A.4}$$

where the **Cartesian covariant metric tensor** $\boldsymbol{\eta_{\mu\nu}}$ is defined as

$$(\eta_{\mu\nu}) \equiv (\eta_{\mu'\nu'}) \equiv \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} . \tag{A0-4}$$

Notice that by following the physicists convention we have set the metric tensor to (+ – – –), while mathematicians generally set it to (– + + +). Modern texts and Wikipedia appear to have settled on the latter convention but I am sticking with the Foster and Nightengale's book, which has the former notation.

From $ds^2 \overset{(A.2)}{=} c^2\, (d\tau)^2$ we see that we can also write equation (A.4) as

$$\boxed{c^2\, (d\tau)^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu = \eta_{\mu'\nu'}\, dx^{\mu'}\, dx^{\nu'}} . \tag{A.5}$$

Equation (A.5) shows that $d\tau$ is invariant across inertial frames. Equation (A.5) is the **special relativity line element**.

Thus, Euclidean spacetime using a Cartesian coordinate system is a pseudo-Riemannian manifold when $g_{\mu\nu} = \eta_{\mu\nu}$ :

Consider a vector where $\lambda^\mu = (\lambda^0, \lambda^1, \lambda^2, \lambda^3) = (1, 1, 1, 1)$. The inner product is

$$\eta_{\mu\nu}\, \lambda^\mu\, \lambda^\nu = \eta_{00}\, \lambda^0\, \lambda^0 + \eta_{11}\, \lambda^1\, \lambda^1 + \eta_{22}\, \lambda^2\, \lambda^2 + \eta_{33}\, \lambda^3\, \lambda^3$$
$$\overset{(A0-4)}{=} 1 - 3 = -2.$$

This shows that $\eta_{\mu\nu}$ is positive indefinite, the criterion for a manifold to be pseudo-Riemannian rather than Riemannian. ✓

This introduces an important dilemma (not mentioned in the book because it does not examine bases in spacetime). In Section 1.10 we made the following definitions for a Riemannian manifold:

$$\mathbf{e}_a \equiv \begin{pmatrix} 0 \\ \vdots \\ 1_a \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_b \equiv \begin{pmatrix} 0 \\ \vdots \\ 1_b \\ \vdots \\ 0 \end{pmatrix}, \quad g_{\text{ab}} = \mathbf{e}_a \bullet \mathbf{e}_b = \delta_{\text{ab}} \quad \Rightarrow \quad \text{matrix } (g_{\text{ab}}) = I\ .$$

Applying this to non-Riemannian spacetime would yield a problem:

$$(g_{\mu\nu}) = (\mathbf{e}_\mu \bullet \mathbf{e}_\nu) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = \eta_{\mu\nu}\ . \qquad \#\#\#\#\#\#$$

To rectify this problem, either the bases need to include imaginary terms, or the definition of dot product must be modified. The approach taken in GR is to leave the basis definition alone, as above, and to revise the **spacetime dot product definition** to match the required result:

$$\mathbf{e}_\mu \bullet \mathbf{e}_\nu \equiv \eta_{\mu\nu} \tag{A0-5}$$

The resulting dot product is then not the Cartesian dot product. The spacetime dot product is an indefinite inner product, re-emphasizing that even though special relativity spacetime is flat, it is a pseudo Euclidean manifold rather than a Euclidean one.

Also observe that if $\boldsymbol{\lambda} = \lambda^\mu \mathbf{e}_\mu$ and $\boldsymbol{\kappa} = \mu^\nu \mathbf{e}_\nu$ are spacetime 4-vectors, then the dot product is the inner product:

$$\boldsymbol{\lambda} \bullet \boldsymbol{\kappa} = \lambda^\mu\, \kappa^\nu\, \mathbf{e}_\mu \bullet \mathbf{e}_\nu = \lambda^\mu \kappa^\nu \eta_{\mu\nu} \tag{A0-6}$$

Recall from Section 1.8 that metric tensors raise and lower indices. This generates **associated spacetime tensors**. For example, $\tau^\mu_\nu = \eta^{\mu\sigma} \tau_{\sigma\nu}$.

Example A.0.1 Let $\lambda^\mu = (\lambda^0, \lambda^1, \lambda^2, \lambda^3)$ and $\kappa^\mu = (\kappa^0, \kappa^1, \kappa^2, \kappa^3)$ be vectors. Show

(a) $\boxed{\lambda_\mu = \left(\lambda^0, -\lambda^1, -\lambda^2, -\lambda^3\right)}$

and in a Cartesian coordinate system the inner product is

(b) $\boxed{\eta_{\mu\nu}\,\lambda^\mu\,\kappa^\nu = \eta^{\mu\nu}\,\lambda_\mu\,\kappa_\nu = \lambda^\mu\,\kappa_\mu = \lambda_\mu\,\kappa^\mu = \lambda^0\,\kappa^0 - \lambda^1\,\kappa^1 - \lambda^2\,\kappa^2 - \lambda^3\,\kappa^3}$ .

Proof.

(1): $\lambda_\mu = \eta_{\mu\nu}\,\lambda^\nu = \eta_{\mu 0}\,\lambda^0 + \eta_{\mu 1}\,\lambda^1 + \eta_{\mu 2}\,\lambda^2 + \eta_{\mu 3}\,\lambda^3 = \begin{cases} \eta_{00}\,\lambda^0 = \lambda^0 & \text{if } \mu = 0 \\ \eta_{\mu\mu}\,\lambda^\mu = -\lambda^\mu & \text{if } \mu > 0 \end{cases}$ . ✓

(2): $\eta_{\mu\nu}\,\lambda^\mu\kappa^\nu = \eta_{00}\,\lambda^0\kappa^0 + \eta_{11}\,\lambda^1\kappa^1 + \eta_{22}\,\lambda^2\kappa^2 + \eta_{33}\,\lambda^3\kappa^3$

$= \lambda^0\kappa^0 - \lambda^1\kappa^1 - \lambda^2\kappa^2 - \lambda^3\kappa^3$ ✓

$\eta_{\mu\nu}\,\lambda^\mu\kappa^\nu = \lambda^\mu\,(\eta_{\mu\nu}\kappa^\nu) = \lambda^\mu\kappa_\mu$ ✓

The other two terms are derived similarly. ■

The **3-space velocity of a particle** is

$$\boldsymbol{v} \equiv \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t}\boldsymbol{i} + \frac{\mathrm{d}y}{\mathrm{d}t}\boldsymbol{j} + \frac{\mathrm{d}z}{\mathrm{d}t}\boldsymbol{k}$$

The **3-space speed of a particle** is defined by $v^2 = \boldsymbol{v} \cdot \boldsymbol{v}$, leading to equations for $v^2$ with respect to inertial frames K and K′:

$$v^2 = \left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2 \quad \text{and} \quad v'^2 = \left(\frac{\mathrm{d}x'}{\mathrm{d}t'}\right)^2 + \left(\frac{\mathrm{d}y'}{\mathrm{d}t'}\right)^2 + \left(\frac{\mathrm{d}z'}{\mathrm{d}t'}\right)^2$$

$$\Leftrightarrow v^2\,\mathrm{d}t^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \quad \text{and} \quad v'^2\mathrm{d}t'^2 = \mathrm{d}x'^2 + \mathrm{d}y'^2 + \mathrm{d}z'^2. \qquad \text{(A0-7)}$$

Thus,

$$c^2\mathrm{d}\tau^2 \overset{\text{(A1-2)}}{=} \mathrm{d}s^2 \overset{\text{(A0-2)}}{=} c^2\,\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \overset{\text{(A0-7)}}{=} c^2\mathrm{d}t^2 - v^2\mathrm{d}t^2 = (c^2 - v^2)\,\mathrm{d}t^2$$

$$\overset{\text{(A0-2)}}{=} (c^2 - v'^2)\,\mathrm{d}t'^2,$$

or,

$$\boxed{\mathrm{d}\tau = \sqrt{1 - \frac{v^2}{c^2}}\;\mathrm{d}t = \sqrt{1 - \frac{v'^2}{c^2}}\;\mathrm{d}t'} . \qquad \text{(A.6)}$$

Consider a particle at rest (i.e., $v = 0$) with respect to an inertial frame K. Equation (A.6) shows that $\mathrm{d}\tau = \mathrm{d}t$, which means that $\tau = t$ + constant, where $t$ is measured by stationary clocks in K. That is, $\tau$ is nothing more than the coordinate time $t$ (up to a constant) for a particle at rest. So,

$$\frac{\mathrm{d}x^0}{\mathrm{d}\tau} = \frac{\mathrm{d}x^0}{\mathrm{d}t} \overset{(A.3)}{=} \frac{\mathrm{d}}{\mathrm{d}t}(ct) = c.$$

This shows that a stationary particle is moving through spacetime at speed $c$ along the $x^0$-axis, which is consistent with equation (A0-3), $c = \frac{\mathrm{d}s}{\mathrm{d}\tau}$, since $s$ is a measure of distance on the $x^0$-axis, its world line.

A non-stationary particle with speed $v > 0$ also travels through spacetime with proper speed $c \overset{(A0-3)}{=} \frac{\mathrm{d}s}{\mathrm{d}\tau}$, and one might expect that some of its speed "bleeds" off into the spatial dimensions. But, curiously, what happens instead is that the spatial speed increases the $x^0$ speed that was previously at speed $c$. This means that the $x^0$ component of proper speed, $\frac{\mathrm{d}\,ct}{\mathrm{d}\tau}$, is greater than $c$ for moving particles:

$$\begin{aligned} c^2 &\overset{(A0-3)}{=} \left(\frac{\mathrm{d}s}{\mathrm{d}\tau}\right)^2 \overset{(A0-2)}{=} \left(\frac{\mathrm{d}\,ct}{\mathrm{d}\tau}\right)^2 - \left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2 - \left(\frac{\mathrm{d}y}{\mathrm{d}\tau}\right)^2 - \left(\frac{\mathrm{d}z}{\mathrm{d}\tau}\right)^2 \\ &= \left(\frac{\mathrm{d}\,ct}{\mathrm{d}\tau}\right)^2 - \left[\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 - \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2 - \left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2\right]\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 \\ &\overset{(A.3)}{=} \left(\frac{\mathrm{d}\,ct}{\mathrm{d}\tau}\right)^2 - \frac{v^2}{\left(1 - \frac{v^2}{c^2}\right)} < \left(\frac{\mathrm{d}\,ct}{\mathrm{d}\tau}\right)^2, \end{aligned}$$

or,

$$\frac{\mathrm{d}\,ct}{\mathrm{d}\tau} > c.$$

## A.1 Lorentz transformations

Definition A **Lorentz transformation** is a coordinate transformation connecting two inertial frames:

$\mathcal{L}: K \to K': \mathcal{L}(x^\mu) = x^{\mu'}$

In matrix form, we could represent $\mathcal{L}$ as

$$x^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu . \qquad \text{(A1-1)}$$

By definition (Section A.0), an inertial frame K′ in flat spacetime moves with respect to inertial frame K at constant velocity and without rotation. (The space portion of K′ may be initially rotated, but when it moves there is no further rotation.) So, $x^{\mu'}$ is obtained from $x^\nu$ by an affine transformation (i.e., linear transformation plus offset)

$$x^{\mu'} = \tilde{\Lambda}^{\mu'}_{\nu} x^\nu + a^\mu \qquad \text{(A.7)}$$

where $\tilde{\Lambda}^{\mu'}_{\nu}$ represents a *linear* transformation and $a^\mu = (0, a^1, a^2, a^3)$ is a constant. Compare to equation (A1-1).

Observe that (0,0,0) is the coordinates of the center O of frame K at $t = 0$, and (0,0,0) is the coordinates of the center O′ of frame K′ at $t' = 0$. So, if $a^\mu = 0$, then the centers O and O′ coincide at $t = t' = 0$, and the Lorentz transformation is called **homogeneous**. Otherwise it is **inhomogeneous**. Some books call the latter transformation a **Poincare transformation** in which case the former is just called a **Lorentz transformation**.

Using this notation, we could call $\tilde{\Lambda}^{\mu'}_{\nu}$ a *homogeneous* transformation matrix whereas $\Lambda^{\mu'}_{\nu}$ is generally inhomogeneous.

The book does not use the tilde notation. Confusingly, it intermingles the two concepts. It uses $\Lambda^{\mu'}_{\nu}$ at times to represent the homogeneous Lorentz transformation and at other times to represent the inhomogeneous transformation. This matters. We show in equation (A1.2), below, that $\tilde{\Lambda}^{\mu'}_{\nu}$ has a very simple expression, but $\Lambda^{\mu'}_{\nu}$ does not.

Caution: In tensor equations, the indices on LHS generally must match the indices on RHS. That would lead to $a^{\mu'}$ on RHS of equation (A.7), which is not correct. Transformation equations are an exception to this rule. Transformation equations have primed terms on one side and unprimed terms on the other side. Despite the way it may look, the expression $\tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu}$ is composed of unprimed terms. This can be seen below in unprimed equation (A1-15), the expansion of $\tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu}$, which is labeled equation (A1-16).

Formula (A.7) can be formally proven using equation (2.32) for how connection coefficients transform:

Since $g_{\mu\nu} = g_{\mu'\,\nu'} = \eta_{\mu\nu} = 0$, 1, or -1, a constant, then $\partial_\sigma g_{\mu\nu} = 0$.

So, $\Gamma^{\alpha}_{\nu\,\sigma} \overset{(2.9)}{=} \Gamma^{\alpha'}_{\nu'\,\sigma'} = 0$ for all $\alpha$, $\nu$, and $\sigma$.

Thus, $0 = \Gamma^{\alpha'}_{\nu'\,\sigma'} \overset{(2.32)}{=} \Gamma^{\mu}_{\beta\,\gamma}\, X^{\alpha'}_{\mu}\, X^{\beta}_{\nu'}\, X^{\gamma}_{\sigma'} + X^{\mu}_{\sigma'\,\nu'}\, X^{\alpha'}_{\mu} = 0 + X^{\mu}_{\sigma'\,\nu'}\, X^{\alpha'}_{\mu} = X^{\mu}_{\sigma'\,\nu'}\, X^{\alpha'}_{\mu}$

$\Rightarrow X^{\mu}_{\sigma'\,\nu'} = 0$ for all $\mu$, $\sigma'$, and $\nu'$.

Swapping primed and unprimed indices gives $X^{\mu'}_{\sigma\,\nu} = 0$ .

Equation (A.7) now follows because if the 2nd derivative of $x^{\mu'}$ is zero, then the first derivative is a constant, and so

$x^{\mu'}$ = (constant) $x^{\nu}$ + (another constant). ✓

More formally, performing the last step using the Fundamental Theorem of calculus, $\int f'(x)\, dx = f(x) + C$:

$$0 = \int 0\, dx^{\sigma} = \int X^{\mu'}_{\sigma\,\nu}\, dx^{\sigma} = \int \frac{\partial}{\partial x^{\sigma}} \left(X^{\mu'}_{\nu}\right) dx^{\sigma} \overset{\text{(Fund Th)}}{=} X^{\mu'}_{\nu} - \tilde{\Lambda}^{\mu'}_{\nu} \quad \Rightarrow \quad X^{\mu'}_{\nu} = \tilde{\Lambda}^{\mu'}_{\nu}$$

$$\Rightarrow \quad x^{\mu'} \overset{\text{(Fund Th)}}{=} \int \frac{\partial}{\partial x^{\nu}} \left(x^{\mu'}\right) dx^{\nu} = \int X^{\mu'}_{\nu}\, dx^{\nu} = \int \tilde{\Lambda}^{\mu'}_{\nu}\, dx^{\nu} \overset{\text{(Fund Th)}}{=} \tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu} + a^{\mu} . \quad \blacksquare$$

As a corollary, we have confirmed that $\tilde{\Lambda}^{\mu'}_{\nu}$ is, in fact, the Jacobian, $X^{\mu'}_{\nu}$. In fact,

$$\tilde{\Lambda}^{\mu'}_{\nu} = X^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^{\nu}}, \quad \text{and} \quad \Lambda^{\mu'}_{\nu} = X^{\mu'}_{\nu} \text{ iff } a^{\mu} = 0: \qquad \text{(A1-2)}$$

$\Lambda^{\mu'}_{\nu}\, x^{\nu} \overset{(A1-1)}{=} x^{\mu'} \overset{(A.7)}{=} \tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu} + a^{\mu}$ for all $x^{\mu'}$

$\Rightarrow \quad \Lambda^{\mu'}_{\nu}\, x^{\nu} = \tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu}$ iff $a^{\mu} = 0$ for all choices of $x^{\nu}$, $\nu = 1 - N$

Choose $x^{\nu} = 0$ for $\nu > 1$. Then $\Lambda^{\mu'}_{\nu}\, x^{\nu} = \Lambda^{\mu'}_{1}\, x^{1}$ and $\tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu} = \tilde{\Lambda}^{\mu'}_{1}\, x^{1}$

$\Rightarrow \quad \Lambda^{\mu'}_{1}\, x^{1} = \tilde{\Lambda}^{\mu'}_{\nu}\, x^{1}$ iff $a^{\mu} = 0$ for all choices of $x^{1}$

$\Rightarrow \quad \Lambda^{\mu'}_{\nu} = \tilde{\Lambda}^{\mu'}_{\nu}$ iff $a^{\mu} = 0$ ✓

We take a small digression to clarify the difference between Lorentz and Jacobian transformations. First, we haven't yet defined, or even used, the term "Jacobian transformation". Rather, we have until now limited our terminology to "Jacobian transformation matrix". However, we showed in Section 0 that every matrix has an associated linear transformation. This association is a natural way to define a Jacobian transformation (even though it is rarely done). We do it now.

Definition A **Jacobian transformation** is a linear transformation $T$ associated with a Jacobian matrix $(X^{\mu'}_{\nu})$.

Since a Jacobian transformation matrix maps a coordinate system at a point P to a coordinate system at a point Q, a Jacobian transformation $T$ is a mapping from the tangent space at P to the tangent space at Q. Recall that a tangent space is not located in any manifold; it is an abstract space associated with a point.

A Lorentz transformation $\mathcal{L}$, however, maps flat spacetime to flat spacetime. Even though we could think of the the tangent "space" at a point in flat spacetime as flat spacetime itself, it is not; it is an associated abstract tangent spacetime, $\mathbf{T}_{\mathrm{P}}$ . So, technically, a Lorentz transformation $\mathcal{L}$ cannot equal a Jacobian transformation $T$ because, as functions, they do not have the same domains and ranges:

$$\mathcal{L} : \mathrm{K} \to \mathrm{K}'$$

$$T : \mathbf{T}_{\mathrm{P}} \to \mathbf{T}_{\mathrm{Q}}$$

Notation In Special and General Relativity, we use $\Lambda^{\mu'}_{\nu}$ to represent (Lorentz) transformations rather than $U^{i'}_{j}$ (Euclidean space) or $X^{a'}_{b}$ (manifolds).

Definition An object $\lambda^{\mu} = (\lambda^0, \boldsymbol{\lambda})$ in spacetime is called a **contravariant vector** if it satisfies the Lorentz transformation formula

$$\lambda^{\mu'} = \Lambda^{\mu'}_{\nu}\, \lambda^{\nu} \qquad \text{(A1-3)}$$

for every pair of inertial frames K and K′, where $\lambda^{\mu'}$ represents the same object but in primed coordinates.

Plugging $dx^{\mu'} = \Lambda^{\mu'}_{\rho} dx^{\rho}$ into equation (A.4):

$$\eta_{\mu\nu}\, dx^{\mu}\, dx^{\nu} \overset{(A.4)}{=} \eta_{\mu'\,\nu'}\, dx^{\mu'} dx^{\nu'} \overset{(A.7)}{=} \eta_{\mu'\,\nu'}\, \Lambda^{\mu'}_{\rho}\, \Lambda^{\nu'}_{\sigma}\, dx^{\rho}\, dx^{\sigma}$$

$$\overset{(\mu\leftrightarrow\rho,\, \nu\leftrightarrow\sigma)}{=} \eta_{\rho'\,\sigma'}\, \Lambda^{\rho'}_{\mu}\, \Lambda^{\sigma'}_{\nu}\, dx^{\mu}\, dx^{\nu}$$

$$\Rightarrow \quad \boxed{\eta_{\mu\nu} = \Lambda^{\rho'}_{\mu}\, \Lambda^{\sigma'}_{\nu}\, \eta_{\rho'\,\sigma'}}\,. \tag{A.8}$$

This is the necessary and sufficient condition for $\eta_{\mu\nu}$ to be a type (0,2) spacetime tensor. This pattern can be extended as was done in Chapter 2 to generate definition (1.72) of a manifold type ($r$, $s$) tensor. The **Lorentz transformation equations for a type (*r*, *s*) spacetime tensor** are:

$$\boxed{\tau^{\mu_1 \ldots \mu_r}_{\nu_1 \ldots \nu_s} = \Lambda^{\mu_1}_{\sigma_1'} \cdots \Lambda^{\mu_r}_{\sigma_r'}\, \Lambda^{\rho_1'}_{\nu_1} \cdots \Lambda^{\rho_s'}_{\nu_s}\, \tau^{\sigma_1' \ldots \sigma_r'}_{\rho_1' \ldots \rho_s'}}\,.$$

Application of this will be made shortly not only for the laws of motion but also for Maxwell's equations of electromagnetism.

While spacetime mechanics is invariant under Lorentz transformations, Newtonian mechanics is not. Rather, Newtonian mechanics is invariant under **Galilean transformations** to which Lorentz transformations reduce when $v\,/\,c$ is negligible. We now expand the development of Lorentz transformations, first for motion along the $x$-axis, followed by motion in 3-dimensions, and then we define and compare the corresponding Galilean transformation.

Definition  A **spacetime boost** (in the x-direction) is an affine transformation such as the following linear (i.e., homogeneous) transformation:

$$\begin{aligned} t' &= B\,t + C\,x \\ x' &= A\,(x - v\,t) \\ y' &= y \\ z' &= z \end{aligned} \tag{A.9}$$

where $A$, $B$, $C \neq 0$. The first two equations in (A.9) represent a linear mixing of space and time due to non-zero velocity in the $x$-direction.

Claim equations (A.9) can be considered to represent the spatial origin $O'$ of $K'$ moving along the $x$-axis of K with a velocity $v$, and the axes coinciding when $t = t' = 0$:

- Along the $x$-axis , $y = z = 0$. Equations (3) and (4) then cause $y' = 0$ and $z' = 0$.
- Equation (2) for the $x'$ component of $O'$ is $0 = A\,(x - v\,t) \;\Leftrightarrow\; x = v\,t$.

Thus, $O'$, and hence all of $K'$, moves in the $+x$ direction with speed $v$. ✓

$$\frac{dx}{dt} = v \qquad \text{(A1-4)}$$

- When $t = t' = 0$, the first equation in (A.9) makes $x = 0$, and then the second equation makes $x' = 0$; that is, the origins coincide. ✓

Thus, we have shown that, as claimed, equations (A.9) describe a homogeneous transformation (i.e., the centers coincide at time $t = t' = 0$). A more general spacetime boost in the $x$-direction would include an offset at time 0.

We can solve for $A$, $B$, and $C$. From equation (A.9) we get

$$\begin{aligned}
&dt' = B\, dt + C\, dx && (dt')^2 = B^2\, dt^2 + 2BC\, dt\, dx + C^2\, dx^2 \\
&dx' = A\,(dx - v\, dt) && (dx')^2 = A^2\left(dx^2 - 2v\, dt\, dx + v^2\, dt^2\right) \\
&dy' = dy && (dy')^2 = dy^2 \\
&dz' = dz && (dz')^2 = dz^2
\end{aligned}$$

Plugging into equation (A0.1) gives

$$c^2\, dt^2 - dx^2 - dy^2 - dz^2 \overset{(A0-2)}{=} c^2\, dt'^2 - dx'^2 - dy'^2 - dz'^2$$
$$= c^2 B^2\, dt^2 + 2c^2 BC\, dt\, dx + c^2 C^2\, dx^2 - A^2\, dx^2 + 2A^2 v\, dt\, dx - A^2 v^2\, dt^2 - dy^2 - dz^2.$$

Comparing coefficients of $dt^2$, $dt\, dx$, and $dx^2$ yields

$$c^2 B^2 - A^2 v^2 = c^2 \iff c^2 B^2 = A^2 v^2 + c^2 \qquad (1)$$
$$c^2 BC + A^2 v = 0 \iff c^2 BC = -A^2 v \qquad (2)$$
$$c^2 C^2 - A^2 = -1 \iff c^2 C^2 = A^2 - 1 \qquad (3)$$

Multiplying equation (1) by (3), and squaring both sides of (2), is solvable, but introduces extraneous solutions that must be discarded:

$$c^4 B^2 C^2 = A^4 v^2 + A^2 c^2 - A^2 v^2 - c^2 \qquad \text{(LHS1 LHS3 = RHS1 RHS3)}$$
$$c^4 B^2 C^2 = A^4 v^2 \qquad \left(\text{LHS2}^2 = \text{RHS2}^2\right)$$

Subtracting yields:

$$0 = A^2 c^2 - A^2 v^2 - c^2 \iff A^2\,(c^2 - v^2) = c^2$$
$$\iff A^2 = \frac{c^2}{c^2 - v^2} = \frac{1}{1 - \frac{v^2}{c^2}} \iff A = \frac{\pm 1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Plugging $A$ into (1):

$$B^2 = \frac{A^2\, v^2 + c^2}{c^2} \quad \Leftrightarrow \quad B = \frac{\pm 1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Plugging $A$ into (3):

$$C^2 = \frac{A^2 - 1}{c^2} \quad \Leftrightarrow \quad C = \frac{\pm \frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

In equations (1) and (3), both “+“ and “-” work for $A$, $B$, and $C$. In equation (2), $A$ can still have either sign, so we choose “+“. Constants $B$ and $C$ must have opposite signs in equation (2), so we are free to choose “+“ for $B$ and “-” for $C$. A solution is thus

$$A = B = \frac{\pm 1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{and} \quad C = \frac{-\frac{v}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \checkmark \tag{A.10}$$

Definition $$\boxed{\gamma \equiv \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{c}{\sqrt{c^2 - v^2}}} \; . \tag{A.11}$$

Observe that $\gamma \geq 1$.

We can choose $A = B = \gamma$ and $C = -\frac{\gamma\, v}{c^2}$, and the boost equations (A.9) becomes:

$$\begin{aligned} t' &= \gamma\,(t - \frac{v}{c^2}\, x) \quad \text{or} \quad c t' = \gamma\, c t - \frac{\gamma\, v}{c}\, x \\ x' &= \gamma\,(x - v\, t) \\ y' &= y \\ z' &= z\,. \end{aligned} \tag{A.12}$$

In matrix form, we write this as

$$\begin{pmatrix} c\,t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{\gamma\, v}{c} & 0 & 0 \\ -\frac{\gamma\, v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\, t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma\, c\, t - x\, \frac{\gamma\, v}{c} \\ -\gamma\, v\, t + x\, \gamma \\ y \\ z \end{pmatrix}. \tag{A.13}$$

Observation 1  When $x = 0$, the second equation gives $x' = -\gamma\, v\, t$. Then the first equation gives $t' = \gamma\, t$, so that $x' = -v\, t'$. This shows that K moves with speed $-v$ with respect to K′, as expected. That is,

$$\frac{\mathrm{dx}'}{\mathrm{d}t'} = -\mathrm{v} \tag{A1-5}$$

Observation 2  $\gamma \geq 1$. From the next observation, this means that $\Delta t > \Delta t'$.

Observation 3  If we take the d$t'$ derivative of both sides of equation 1 of (A.12) we get

$$1 = \frac{\mathrm{d}t'}{\mathrm{d}t'} = \gamma\left(\frac{\mathrm{d}t}{\mathrm{d}t'} - \frac{v}{c^2}\frac{\mathrm{dx}}{\mathrm{d}t'}\right) = \gamma\frac{\mathrm{d}t}{\mathrm{d}t'}\left[1 - \left(\frac{v}{c^2}\frac{\mathrm{dx}}{\mathrm{d}t'}\right)\frac{\mathrm{d}t'}{\mathrm{d}t}\right] = \gamma\frac{\mathrm{d}t}{\mathrm{d}t'}\left(1 - \frac{v}{c^2}\frac{\mathrm{dx}}{\mathrm{d}t}\right)$$

$$\overset{(A1-4)}{=} \gamma\frac{\mathrm{d}t}{\mathrm{d}t'}\left(1 - \frac{v^2}{c^2}\right) = \gamma\frac{\mathrm{d}t}{\mathrm{d}t'}\frac{1}{\gamma^2} = \frac{1}{\gamma}\frac{\mathrm{d}t}{\mathrm{d}t'}$$

$$\Rightarrow \quad \boxed{\frac{\mathrm{d}t}{\mathrm{d}t'} = \gamma} \tag{A1-6}$$

The book's development of spacetime boost is limited to the $x$-direction. The approach generalizes to a boost in an arbitrary space direction, developed now.

Definition  The spacetime boost (A.12), above, is called the **1+1 homogeneous Lorentz transformation with speed** $v$. The **3+1 Lorentz homogeneous transformation with velocity** $\boldsymbol{v}$ is a transformation where motion is in an arbitrary space direction. It has the equation

$$\begin{aligned} t' &= B\, t + \boldsymbol{C} \cdot \boldsymbol{x} \\ \boldsymbol{x}' &= A\,(\boldsymbol{x} - \boldsymbol{v}\, t) \end{aligned} \tag{A1-7}$$

where we denote coordinates $x^\mu = (ct, x, y, z) = (x^0, \boldsymbol{x})$. We use boldface to denote $\boldsymbol{x} = (x, y, z)$ and $\boldsymbol{v} = (v_x, v_y, v_z)$. Note that $A$ and $B$ must be scalars and $\boldsymbol{C}$ must be a 3-vector in order that the RHS of the first equation be a scalar and the RHS of the second equation be a space 3-vector.

Observe that when $\boldsymbol{x}' = 0$, then $\boldsymbol{x} = \boldsymbol{v}\, t$, which shows that the center O′, and hence all of frame K′, moves with 3-space velocity $\boldsymbol{v}$. When $t = t' = 0$, the first equation shows that $\boldsymbol{x} = 0$ and the second equation shows that $\boldsymbol{x}' = 0$. Thus, Equations (A1-7) describe an inertial frame K′ having velocity $\boldsymbol{v}$ whose center O′ coincides with center O at time zero, confirming that equations (A.17) describe a *homogeneous* Lorentz transformation.

As in the 1+1 case, we can solve for $A$, $B$, and $\boldsymbol{C}$ by using the equation

$$c^2\,(\mathrm{d}t')^2 - (\mathrm{d}\boldsymbol{x}')^2 = c^2\,(\mathrm{d}t)^2 - (\mathrm{d}\boldsymbol{x})^2. \tag{A1-8}$$

$$\mathrm{d}t' = B\,\mathrm{d}t + \boldsymbol{C}\bullet\mathrm{d}\boldsymbol{x} \qquad (\mathrm{d}t')^2 = B^2\,\mathrm{d}t^2 + 2B\boldsymbol{C}\,\mathrm{d}t\bullet\mathrm{d}\boldsymbol{x} + (\boldsymbol{C}\bullet\mathrm{d}\mathbf{x})^2$$

$$\mathrm{d}\boldsymbol{x}' = A\,(\mathrm{d}\boldsymbol{x} - \boldsymbol{v}\,\mathrm{d}t) \qquad (\mathrm{d}x')^2 = A^2\left(\mathrm{d}x^2 - 2\boldsymbol{v}\,\mathrm{d}t\bullet\mathrm{d}\boldsymbol{x} + v^2\,\mathrm{d}t^2\right)$$

Plugging into equation (A0.1) gives

$$c^2\,(\mathrm{d}t)^2 - (\mathrm{d}x)^2 \overset{(A0-2)}{=} c^2\,(\mathrm{d}t')^2 - (\mathrm{d}x')^2$$

$$= c^2B^2\,\mathrm{d}t^2 + 2c^2B\boldsymbol{C}\bullet\mathrm{d}t\,\mathrm{d}\mathbf{x} + c^2\,(\boldsymbol{C}\bullet\mathrm{d}\mathbf{x})^2 - A^2\left(\mathrm{d}x^2 - 2\boldsymbol{v}\bullet\mathrm{d}t\,\mathrm{d}\boldsymbol{x} + v^2\,\mathrm{d}t^2\right)$$

Comparing coefficients of $(\mathrm{d}t)^2$ and $(\bullet\,\mathrm{d}t\,\mathrm{d}\mathbf{x})$ yields

$$c^2B^2 - A^2v^2 = c^2 \iff c^2B^2 = A^2v^2 + c^2 \tag{1}$$

$$c^2B\boldsymbol{C} + A^2v = 0 \iff c^2B\boldsymbol{C} = -A^2v \tag{2}$$

There is a 3rd equation, a vector equation, for the coefficients of $(\mathrm{d}x)^2$. It can be broken down into three pieces: $(\mathrm{d}x^1)^2$, $(\mathrm{d}x^2)^2$, and $(\mathrm{d}x^3)^2$, and is complicated to express. Nonetheless, in principle, this process generates 5 equations in 5 unknowns, the unknowns being $A$, $B$, and $C^1$, $C^2$, and $C^3$. In principle, we can generate the remaining three equations, and solve the system of equations. In practice, it is easier to guess the solution and then work backwards to check it. Modeled after the 1+1 solution, we guess that

$$A = B = \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{c}{\sqrt{c^2-v^2}} \quad \text{and} \quad \boldsymbol{C} = -\frac{\gamma}{c^2}\boldsymbol{v}\,. \tag{A1-9}$$

To confirm this solution, we show that equation (A1-8) holds.

Since points in the K′ frame move with velocity $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v}$,

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{v}\,\mathrm{d}t\,, \qquad \boldsymbol{v}\bullet\mathrm{d}\boldsymbol{x} = v^2\,\mathrm{d}t\,, \qquad \text{and} \qquad (\mathrm{d}\boldsymbol{x})^2 = \mathrm{d}\boldsymbol{x}\bullet\mathrm{d}\boldsymbol{x} = v^2\,\mathrm{d}t^2\,. \tag{A1-10}$$

Relative to the K′ frame, points $\boldsymbol{x}'$ do not move:

$$\frac{\mathrm{d}\boldsymbol{x}'}{\mathrm{d}t'} = 0\,. \tag{A1-11}$$

By differentiating equations (A1-7) we get

$$\mathrm{d}t' = B\,\mathrm{d}t + \boldsymbol{C}\bullet\mathrm{d}\boldsymbol{x}$$
$$\mathrm{d}\boldsymbol{x}' = A\,[\mathrm{d}\boldsymbol{x} - \boldsymbol{v}\,\mathrm{d}t\,] \qquad \text{(A1-12)}$$

$$\Rightarrow \quad (\mathrm{d}t')^2 = B^2\,(\mathrm{d}t)^2 + 2\,B\,\boldsymbol{C}\bullet\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + (\boldsymbol{C}\bullet\mathrm{d}\boldsymbol{x})^2$$
$$\overset{\text{(A1-9)}}{=} \gamma^2\,(\mathrm{d}t)^2 + 2\gamma\,(-\frac{\gamma}{c^2}\boldsymbol{v})\bullet\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + (-\frac{\gamma}{c^2}\boldsymbol{v}\bullet\mathrm{d}\boldsymbol{x})^2$$
$$= \gamma^2\,(\mathrm{d}t)^2 - \frac{2\,\gamma^2}{c^2}\boldsymbol{v}\bullet\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + \frac{\gamma^2}{c^4}(\boldsymbol{v}\bullet\mathrm{d}\boldsymbol{x})^2$$
$$\overset{\text{(A1-10)}}{=} \gamma^2\,(\mathrm{d}t)^2 - \frac{2\,\gamma^2}{c^2}v^2\,\mathrm{d}t^2 + \frac{\gamma^2}{c^4}v^4\,\mathrm{d}t^2$$
$$= \gamma^2\,(\mathrm{d}t)^2\,(1 - 2\frac{v^2}{c^2} + \frac{v^4}{c^4}) = \gamma^2\,(\mathrm{d}t)^2\,(1 - \frac{v^2}{c^2})^2$$
$$\Rightarrow \quad \mathrm{d}t' = \gamma\,\mathrm{d}t\,(1 - \frac{v^2}{c^2}) \overset{(A.11)}{=} \gamma\,\mathrm{d}t\,\frac{1}{\gamma^2} = \frac{\mathrm{d}t}{\gamma}$$
$$\Rightarrow \quad \boxed{\frac{\mathrm{d}t}{\mathrm{d}t'} = \gamma} \qquad \text{(A1-13)}$$

As expected, this agrees with the 1+1 case, equation (A1-6).

Also, from (A1-12), we get

$$(\mathrm{d}\boldsymbol{x}')^2 = 0 : \qquad \text{(A1-14)}$$

$$(\mathrm{d}\boldsymbol{x}')^2 = A^2\,[\,(\mathrm{d}\boldsymbol{x})^2 - 2\,\boldsymbol{v}\bullet\mathrm{d}\boldsymbol{x}\,\mathrm{d}t + v^2\,\mathrm{d}t^2\,]$$
$$\overset{\text{(A1-10)}}{=} \gamma^2\,[\,v^2 - 2v^2 + v^2\,]\,\mathrm{d}t^2 = 0. \qquad ✓$$

Though perhaps a little strange at first, this is consistent with (A1-11) that $\frac{\mathrm{d}\boldsymbol{x}'}{\mathrm{d}t'} = 0$ .

We are now ready to show our guesses for $A$, $B$, and $\boldsymbol{C}$ are correct, that equation (A1-8) holds:

$$c^2\,(\mathrm{d}t')^2 - (\mathrm{d}\boldsymbol{x}')^{\,2} \overset{\text{(A1-13, A1-14)}}{=} c^2\,\frac{(\mathrm{d}t)^2}{\gamma^2} - 0 = c^2\,\frac{c^2 - v^2}{c^2}(\mathrm{d}t)^2 = (c^2 - v^2)\,(\mathrm{d}t)^2$$
$$= c^2\,(\mathrm{d}t)^2 - v^2\,\mathrm{d}t^2 \overset{\text{(A1-10)}}{=} c^2\,(\mathrm{d}t)^2 - (\mathrm{d}\boldsymbol{x})^2. \qquad ✓$$

Having established values for $A$, $B$, and $\boldsymbol{C}$ in (A1-9), we plug these values into the 3+1 Lorentz transformation (A1-7) to generate

$$t' = \gamma\,(t\; - \frac{\boldsymbol{v} \bullet \boldsymbol{x}}{c^2}) \qquad \text{or} \qquad ct' = \gamma\,(ct\; - \frac{\boldsymbol{v}}{c} \bullet \boldsymbol{x})$$
$$\boldsymbol{x}' = \gamma\,(\boldsymbol{x} - \boldsymbol{v}\, t) \tag{A1-15}$$

and, in matrix form,

$$x^{\mu'} = \tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu} : \tag{A1-16}$$

$$\begin{pmatrix} c\,t' \\ \\ \boldsymbol{x}' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\frac{\boldsymbol{v}^{\mathrm{T}}}{c} \\ -\frac{\boldsymbol{v}}{c} & \mathrm{I} \end{pmatrix} \begin{pmatrix} c\,t \\ \\ \boldsymbol{x} \end{pmatrix} = \gamma \begin{pmatrix} c\,t - \frac{\boldsymbol{v} \bullet \boldsymbol{x}}{c} \\ \\ -\boldsymbol{v}\,t + \boldsymbol{x} \end{pmatrix}, \tag{A1-17}$$

where $\tilde{\Lambda}^{\mu'}_{\nu}$ is the *homogeneous* Lorentz matrix, $\boldsymbol{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$, $\boldsymbol{v}^T = (v_x\;\; v_y\;\; v_z)$, and $\mathrm{I}$ is the 3x3 identity matrix. We use the transpose of $\boldsymbol{v}$ in the (1,2) element of the matrix because the row vector $\boldsymbol{v}^{\mathrm{T}}$ times the column vector $\boldsymbol{x}$ yields the dot product $\boldsymbol{v} \bullet \boldsymbol{x}$. This completes the development of the (3+1) boost. ■

As in the 1+1 case, we observe that when $\boldsymbol{x} = 0$, the 2nd equation gives $\boldsymbol{x}' = \gamma\, \boldsymbol{v}\, t$. Then the first equation gives $t' = \gamma\, t$, so that $\boldsymbol{x}' = -\,\boldsymbol{v}\, t'$. This confirms that K moves with velocity $-\,\boldsymbol{v}$ with respect to K′; i.e., $\boldsymbol{v}' = -\,\boldsymbol{v}$.

Using tilde (~) for the homogeneous case, the matrix

$$\tilde{\Lambda} \equiv \gamma \begin{pmatrix} 1 & -\frac{\boldsymbol{v}^{\mathrm{T}}}{c} \\ -\frac{\boldsymbol{v}}{c} & \mathrm{I} \end{pmatrix} \tag{A1-18}$$

is the general form for the homogeneous Lorentz transformation matrix. This can also be expressed in expanded form:

$$\tilde{\Lambda} = \gamma \begin{pmatrix} 1 & -\frac{v_x}{c} & -\frac{v_y}{c} & -\frac{v_z}{c} \\ -\frac{v_x}{c} & 1 & 0 & 0 \\ -\frac{v_y}{c} & 0 & 1 & 0 \\ -\frac{v_z}{c} & 0 & 0 & 1 \end{pmatrix}. \tag{A1-19}$$

Equations (A1-15) and (A1-19) can be used to provide a computational verification of equation (A1-2) that $\tilde{\Lambda}^{\mu'}_{\nu} = \frac{\partial x^{\mu'}}{\partial x^{\nu}}$:

Since $\boldsymbol{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\boldsymbol{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$, we can break out $x^{i'}$ from equation (A1-15):

$$x^{i'} = \gamma\left(x^i - v^i t\right).$$

For example,

$$x^{1'} = \gamma\,(x - v_x t).$$

Next,

$$\frac{\partial x^{o'}}{\partial x^0} \overset{(A1-15)}{=} \gamma\,\Big(\frac{\partial(ct)}{\partial(ct)} - \frac{\boldsymbol{v}}{c} \bullet \frac{\partial \boldsymbol{x}}{\partial(ct)}\Big) = \gamma + 0 = \gamma \overset{(A1-18)}{=} \tilde{\Lambda}^{0'}_{0} \qquad \checkmark$$

$$\frac{\partial x^{o'}}{\partial x^1} = \gamma\,\Big(\frac{\partial(ct)}{\partial x} - \frac{\boldsymbol{v}}{c} \bullet \frac{\partial \boldsymbol{x}}{\partial x}\Big) = \gamma\,\Big(0 - \frac{1}{c}\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\Big) = -\gamma\,\frac{v_x}{c} \overset{(A1-19)}{=} \tilde{\Lambda}^{0'}_{1} \qquad \checkmark$$

$$\frac{\partial x^{1'}}{\partial x^0} = \gamma\,\Big(\frac{\partial x}{\partial(ct)} - v_x\,\frac{\partial t}{\partial(ct)}\Big) = \gamma\,\frac{v_x}{c} \overset{(A1-19)}{=} \tilde{\Lambda}^{1'}_{0} \qquad \checkmark$$

$$\frac{\partial x^{1'}}{\partial x^1} = \gamma\,\Big(\frac{\partial x}{\partial x} - v_x\,\frac{\partial t}{\partial x}\Big) = \gamma \overset{(A1-19)}{=} \tilde{\Lambda}^{1'}_{1} \qquad \checkmark$$

Etc. $\checkmark$

The general (non-homogeneous) Lorentz transformation matrix Λ in equation (A1-20), proven below, does not equal $\tilde{\Lambda}$ as given by equation (A1-19), and so we conclude that $\Lambda \neq \frac{\partial x^{\mu'}}{\partial x^{\nu}}$. When we compute changes of coordinates, we either need to make a homogeneous change of coordinates or else use the messy matrix (A1-20) or (A1-21).

The general Lorentz transformation matrix is given by Frobenius at StackExchange.com (search for Transformation of 4–velocity) as

$$\boxed{\Lambda = \begin{pmatrix} \gamma & -\frac{\gamma \boldsymbol{v}^{\mathrm{T}}}{c} \\ -\frac{\gamma \boldsymbol{v}}{c} & \mathrm{I} + \frac{\gamma^2}{c^2(\gamma+1)} \mathbf{v}\mathbf{v}^{\mathrm{T}} \end{pmatrix}} \tag{A1-20}$$

which equals

$$\Lambda = \begin{pmatrix} \gamma & -\frac{\gamma v_x}{c} & -\frac{\gamma v_y}{c} & -\frac{\gamma v_z}{c} \\ -\frac{\gamma v_x}{c} & 1+(\gamma-1)\frac{v_x{}^2}{v^2} & (\gamma-1)\frac{v_x v_y}{v^2} & (\gamma-1)\frac{v_x v_z}{v^2} \\ -\frac{\gamma v_y}{c} & (\gamma-1)\frac{v_y v_x}{v^2} & 1+(\gamma-1)\frac{v_y{}^2}{v^2} & (\gamma-1)\frac{v_y v_z}{v^2} \\ -\frac{\gamma v_z}{c} & (\gamma-1)\frac{v_z v_x}{v^2} & (\gamma-1)\frac{v_z v_y}{v^2} & 1+(\gamma-1)\frac{v_z{}^2}{v^2} \end{pmatrix} \tag{A1-21}$$

from Wikipedia (search for Lorentz transformation matrix). We easily see that these matrices are equal by using $\frac{\gamma-1}{v^2} = \frac{\gamma^2}{c^2(\gamma+1)}$ from equation (A1-23), below, with the understanding that a column vector $\boldsymbol{v}$ times a row vector $\boldsymbol{v}^{\mathrm{T}}$ is

$$\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} (\, v_x \;\; v_y \;\; v_z \,) = \begin{pmatrix} v_x{}^2 & v_x v_y & v_x v_z \\ v_y v_x & v_y{}^2 & v_y v_z \\ v_z v_x & v_z v_y & v_z{}^2 \end{pmatrix}.$$

The usual method for generating the matrix $\Lambda$ involves a somewhat complex process of combining a sequence of transformations that includes a 3-space rotation, a spatial translation (i.e., an offset), and a boost in an arbitrary direction. However, we can derive it quite easily from the homogeneous Lorentz matrix $\tilde{\Lambda}$.

Using $\Lambda$ from equation (A1-20), consider the equations $x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu}$:

$$t' = \gamma\,(t - \frac{\boldsymbol{v} \bullet \boldsymbol{x}}{c^2}) \quad \text{or} \quad ct' = \gamma\,(ct - \frac{\boldsymbol{v} \bullet \boldsymbol{x}}{c})$$

$$\boldsymbol{x}' = \boldsymbol{x} + \frac{\gamma^2}{c^2(\gamma+1)} (\boldsymbol{v} \bullet \boldsymbol{x})\, \boldsymbol{v} - \frac{\gamma \boldsymbol{v}}{c} c\, t \tag{A1-22}$$

Claim these equations equal the homogeneous equations (A1-16) plus an offset, $a^\mu$:

The expressions for $t'$ are identical (and the offset time component is 0). ✓

To see that the expressions for $\boldsymbol{x}'$ only differ by an offset, we first develop the equation for d$\boldsymbol{x}'$ :

$$\gamma = 1 + \frac{v^2}{c^2}\frac{\gamma^2}{\gamma+1} : \tag{A1-23}$$

$$\gamma^2 = \frac{c^2}{c^2 - v^2} \quad \Leftrightarrow \quad c^2\gamma^2 - v^2\gamma^2 = c^2$$

$$\Leftrightarrow \quad c^2\gamma^2 = c^2 + v^2\gamma^2 \tag{A1-24}$$

$$\Rightarrow \quad \gamma\, c^2(\gamma + 1) = c^2\gamma + c^2\gamma^2 \overset{(A1-16)}{=} c^2\gamma + c^2 + v^2\gamma^2 = c^2(\gamma + 1) + v^2\gamma^2.$$

Dividing both sides by $c^2(\gamma + 1)$ gives $\gamma = 1 + \frac{v^2}{c^2}\frac{\gamma^2}{\gamma+1}$ ✓

$$\gamma\, \mathrm{d}\boldsymbol{x} = \mathrm{d}\boldsymbol{x} + \frac{\gamma^2}{c^2\,(\gamma+1)}\,(\boldsymbol{v} \bullet d\boldsymbol{x})\,\boldsymbol{v} : \tag{A1-25}$$

$$\gamma\, \mathrm{d}\boldsymbol{x} \overset{(A1-23)}{=} \mathrm{d}\boldsymbol{x} + \frac{v^2}{c^2}\frac{\gamma^2}{\gamma+1}\,\mathrm{d}\boldsymbol{x} \overset{(A1-10)}{=} \mathrm{d}\boldsymbol{x} + \frac{\gamma^2}{c^2\,(\gamma+1)}\,(v^2\,\mathrm{d}t)\,\mathbf{v}$$

$$\overset{(A1-10)}{=} \mathrm{d}\boldsymbol{x} + \frac{\gamma^2}{c^2\,(\gamma+1)}\,(\boldsymbol{v} \bullet \mathrm{d}\boldsymbol{x})\,\boldsymbol{v} \qquad ✓$$

$$\mathrm{d}\boldsymbol{x}' \overset{(A1-15)}{=} \gamma\, \mathrm{d}\boldsymbol{x} - \gamma\, \boldsymbol{v}\, \mathrm{d}t \overset{(A1-25)}{=} \mathrm{d}\boldsymbol{x} + \frac{\gamma^2}{c^2(\gamma+1)}\,(\boldsymbol{v} \bullet \mathrm{d}\boldsymbol{x})\,\boldsymbol{v} - \frac{\gamma\,\boldsymbol{v}}{c}\,c\,\mathrm{d}t \qquad ✓$$

We integrate the right three terms and then the middle two terms to get

$$\boldsymbol{x} + \frac{\gamma^2}{c^2\,(\gamma+1)}\,(\boldsymbol{v} \bullet \boldsymbol{x})\,\boldsymbol{v} - \frac{\gamma\,\boldsymbol{v}}{c}\,c\,t \;= \gamma\,(\boldsymbol{x} - \boldsymbol{v}\,t) + \boldsymbol{a} \tag{A1-26}$$

where $\boldsymbol{a}$ is the spatial offset.

So, the spacetime offset, composed of the time and space offsets, is $a^\mu = \begin{pmatrix} 0 \\ \boldsymbol{a} \end{pmatrix}$.

Putting this altogether, we confirm the claim, essentially by combining equations (A1-1) and (A.7):

$$\boxed{x^{\mu'} = \Lambda^{\mu'}_{\nu}\, x^{\nu} = \tilde{\Lambda}^{\mu'}_{\nu}\, x^{\nu} + a^{\mu}} : \tag{A1-27}$$

$$x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu} = \begin{pmatrix} \gamma & -\frac{\gamma \mathbf{v}^{\mathrm{T}}}{c} \\ -\frac{\gamma \mathbf{v}}{c} & \mathrm{I} + \frac{\gamma^2}{c^2(\gamma^2+1)} \mathbf{v}\mathbf{v}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} c\,t \\ \boldsymbol{x} \end{pmatrix}$$

$$= \begin{pmatrix} \gamma\left(ct - \frac{\boldsymbol{v} \bullet \boldsymbol{x}}{c}\right) \\ -\gamma\, \boldsymbol{v}\, t + \boldsymbol{x} + \frac{\gamma^2}{c^2(\gamma^2+1)} (\boldsymbol{v} \bullet \boldsymbol{x})\, \boldsymbol{v} \end{pmatrix} \overset{(\text{A1}-26)}{=} \begin{pmatrix} \gamma\left(ct - \frac{\boldsymbol{v} \bullet \boldsymbol{x}}{c}\right) \\ \gamma\,(\boldsymbol{x} - \boldsymbol{v}\, t) + \boldsymbol{a} \end{pmatrix}$$

$$= \gamma \begin{pmatrix} 1 & -\frac{\boldsymbol{v}^{\mathrm{T}}}{c} \\ -\frac{\boldsymbol{v}}{c} & \mathrm{I} \end{pmatrix} \begin{pmatrix} c\,t \\ \boldsymbol{x} \end{pmatrix} + \begin{pmatrix} 0 \\ \boldsymbol{a} \end{pmatrix} = \tilde{\Lambda}^{\mu'}_{\nu} x^{\nu} + a^{\mu} \qquad \checkmark$$

That is, the general Lorentz transformation can be expressed as the homogeneous Lorentz transformation plus an offset. I have never seen the homogeneous matrix referenced. However, it is easier to use and can be appropriate in many situations. For example, in Section A.7 I use it to derive the Doppler frequency shift, which is not affected by translation to an offset, and thus only needs the simpler homogeneous transformation.

Definition $\boldsymbol{\psi}$: $\quad \tanh \psi \equiv \frac{v}{c}.$ (A1-28)

Claim: $\quad \gamma = \cosh \psi$ : (A1-29)

Since $1 - \tanh^2 \psi = \operatorname{sech}^2 \psi$, then

$$\gamma \overset{(\text{A}.11,\ \text{A1}-29)}{=} \frac{1}{\sqrt{1 - \tanh^2 \psi}} = \frac{1}{\operatorname{sech} \psi} = \cosh \psi \qquad \checkmark$$

$$\frac{v}{c} \cosh \psi \overset{(\text{A1}-28)}{=} \tanh \psi \cosh \psi = \sinh \psi , \tag{A1-30}$$

$$c\,t' \overset{(A.12)}{=} c\gamma\,(t - \frac{v}{c^2}x)$$
$$\overset{(A1-29)}{=} c\,t\cosh\psi - x\,\frac{v}{c}\cosh\psi \overset{(A1-30)}{=} c\,t\cosh\psi - x\sinh\psi\,, \tag{A1-31}$$
$$x' \overset{(A.12)}{=} \gamma\,(x - v\,t)$$
$$\overset{(A1-29)}{=} x\cosh\psi - c\,t\,\frac{v}{c}\cosh\psi \overset{(A1-30)}{=} -c\,t\sinh\psi + x\cosh\psi\,, \tag{A1-32}$$

and we can express the $x$-boost, equation (A.12), as

$$\begin{aligned} c\,t' &\overset{(A1-31)}{=} c\,t\cosh\psi - x\sinh\psi \\ x' &\overset{(A1-32)}{=} -c\,t\sinh\psi + x\cosh\psi \\ y' &= y \\ z' &= z\,. \end{aligned} \tag{A.14}$$

In matrix form this is

$$\begin{pmatrix} c\,t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0 \\ -\sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\,t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c\,t\cosh\psi - x\sinh\psi \\ -c\,t\sinh\psi + x\cosh\psi \\ y \\ z \end{pmatrix}. \tag{A1-33}$$

Definition A **Galilean boost** (in the $x$-direction) is a transformation

$$\begin{aligned} t' &= t \\ x' &= x - v\,t \\ y' &= y \\ z' &= z \end{aligned} \tag{A1-34}$$

Exercise A.1.3 A Lorentz boost, equations (A.12), reduces to a Galilean boost when $v << c$. This is because when $v << c$, $\gamma \overset{(A.11)}{\approx} 1 \Rightarrow t' \overset{(A.12)}{\approx} t$ and $x' \overset{(A.12)}{\approx} x - v\,t$ ✓

A **spacetime rotation generated by spinning the *xy*-plane** can be described by the rotation matrix

$$\begin{pmatrix} c\,t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\,t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ x\cos\theta - y\sin\theta \\ x\sin\theta + x\cos\theta \\ z \end{pmatrix}. \tag{A1-35}$$

Observe that the spacetime boost matrix in equation (A1-33) somewhat resembles the rotation matrix in equation (A1-35). A graphical comparison between the boost and the rotation is possible and is illustrated in Figure A.1, below.

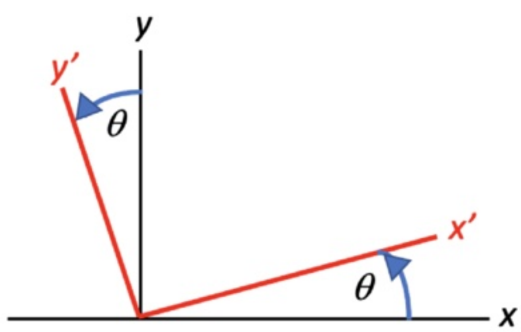



Figure A.1.a

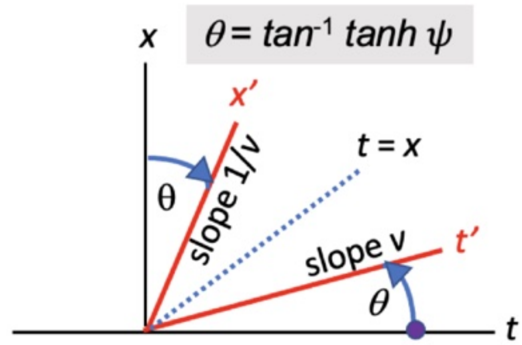



Figure A.1.b

The LHS represents the spacetime rotation; the RHS represents the spacetime boost. For simplicity, we assume units where $c$ = 1 so that we can label the time-axis by $t$ rather than $ct$. Also, since $|v| \leq c$, this means that $|v| \leq 1$.

On RHS, the $t'$-axis has equation $x'$ = 0, which by equation (A.12) yields

$$x - v\,t = 0 \;\Leftrightarrow\; x = v\,t \overset{(A1-24)}{=} t \tanh \psi\,.$$

So, the slope of the $t'$-axis is $\dfrac{\Delta x}{\Delta t}$ = $v$. If we fix $\psi$ and define a fixed $\theta$ by $\tan\theta = \tanh\psi$, then $x = t \tanh \psi = t \tan\theta$. This is shown in Figure A.1.b where the

the $t'$-axis has the equation $x = t \tan\theta$ and slope $v$.

The $x'$-axis in Figure A.1.b has equation $t'$ = 0, which by equation (A.12) yields

$$t - v\,x = 0 \;\Leftrightarrow\; x = \frac{1}{v}\,t \overset{(A1-24)}{=} \frac{t}{\tanh \psi}\,.$$

In Figure A.1.b,

the $x'$-axis has equation $x = \dfrac{t}{\tan\theta}$ and slope $1/v$.

A rotation rotates two axes in the same direction. We see that a spacetime boost rotates the $x$ and $t$ axes in opposite directions.

In the limit as $v \to 1$,

the slope of the $t'$-axis = $v \to 1$ from below

the slope of the $t'$-axis = $\dfrac{1}{v} \to 1$ from above

The line $x = t$ represents a particle traveling at the speed of light in the $x$ direction.

Spatial rotations about an axis (whether the $x$, $y$, or $z$ axes, or a skew axis) are not Lorentz transformation, but they *are* homogeneous linear transformations. All of spacetime gets rotated, but the $t$-axis stays fixed. In spacetime, if we try to “spin”, say, the $xt$-plane, we mix space and time, and so we get a boost, not a rotation. Any rotation of all of spacetime leaves the time axis alone. So, strangely, a spacetime rotation has two fixed axes, a spatial axis and the $t$-axis. This 4-dimensional quirk is difficult to visualize because we live in 3 dimensions.

## A.2 Relativistic addition of velocities

Suppose we have three inertial frames, K, K′, and K′′ such that K′ represents a boost having speed $v$ in the $x$ direction with respect to K, and K′′ represents a boost having speed $w$ with respect to K′ in the $x'$direction. Equations (A.14) hold for K and K′:

$$\begin{aligned} c\,t' &\overset{(A1-31)}{=} c\,t\cosh\psi - x\sinh\psi \\ x' &\overset{(A1-32)}{=} -c\,t\sinh\psi + x\cosh\psi \\ y' &= y \\ z' &= z \end{aligned} \tag{A.14}$$

where $\psi$ is defined by equation (A1-24): $\tanh\psi \equiv \frac{v}{c}$.

Analogously, the K′ boost equations involving K′ and K′′ are

$$\begin{aligned} c\,t'' &= c\,t'\cosh\phi - x'\sinh\phi \\ x'' &= -c\,t'\sinh\phi + x'\cosh\phi \\ y'' &= y' \\ z'' &= x' \end{aligned} \tag{A.15}$$

with $\phi$ defined by $\tanh\phi = \frac{v'}{c}$.

Substituting $c\,t'$, $x'$, $y'$, and $z'$ from equations (A.14) yields

$$\begin{aligned} c\,t'' &= c\,t\cosh(\psi+\phi) - x\sinh(\psi+\phi) \\ x'' &= -c\,t\sinh(\psi+\phi) + x\cosh(\psi+\phi) \\ y'' &= y \\ z'' &= x \end{aligned} \tag{A.16}$$

because $\cosh(\psi+\phi) = \cosh\psi\cosh\phi + \sinh\psi\sinh\phi$

$\sinh(\psi+\phi) = \sinh\psi\cosh\phi + \cosh\psi\sinh\phi$.

This shows that K$''$ represents a boost having speed *u* in the *x* direction of K, where

$$\boxed{u = \frac{v + w}{1 + \frac{v\,w}{c^2}}} : \tag{A.17}$$

$$\begin{aligned}
\tanh(\psi + \phi) &= \frac{\sinh\psi\cosh\phi + \cosh\psi\sinh\phi}{\cosh\psi\cosh\phi + \sinh\psi\sinh\phi} \\
&= \frac{\frac{\sinh\psi\cosh\phi}{\cosh\psi\cosh\phi} + \frac{\cosh\psi\sinh\phi}{\cosh\psi\cosh\phi}}{\frac{\cosh\psi\cosh\phi}{\cosh\psi\cosh\phi} + \frac{\sinh\psi\sinh\phi}{\cosh\psi\cosh\phi}} \\
&= \frac{\tanh\psi + \tanh\phi}{1 + \tanh\psi\tanh\phi}
\end{aligned}$$

$$u \overset{(A1-24)}{\equiv} c\tanh(\psi + \phi) = c\,\frac{\tanh\psi + \tanh\phi}{1 + \tanh\psi\tanh\phi} \overset{(A1-24)}{=} c\,\frac{\frac{v}{c} + \frac{w}{c}}{1 + \frac{v}{c}\frac{w}{c}} = \frac{v + w}{1 + \frac{v\,w}{c^2}} \quad \checkmark$$

Formula (A.17) is the **relativistic formula for addition of velocities**, replacing the Newtonian formula $u = v + w$. Exercise A.2.2 proves that $u < c$.

## A.3 Simultaneity

In Newtonian physics, events occur at the same time in a moving frame as in a stationary frame. That is, time is absolute. All frames experience the same time.

This is not the case in relativity. Consider a ball tossed upward and then caught by a person on a train moving a speed $v$ along the $x$-axis. Suppose a stationary observer watches the train. Let K be the frame of the observer and K′ that of the train. Assume the ball is tossed at time 0 in both frames. Suppose the observer sees the ball land at time $t_0$ at location $x_0$ in his frame K. From boost equation (A.12), the time $t_0'$ in frame K′ that the ball lands is $t_0' = \gamma\,(t_0 - \frac{v}{c^2}\,x_0)$, different than time $t_0$ because it depends also on $x_0$. (Also, see Figure A.5).

## A.4 Time dilation and length contraction

Equation (A.6) gives the proper time interval $\Delta\tau$ recorded by a clock moving with speed $v$ < c relative to an inertial frame K:

$$\boxed{\Delta\tau = \sqrt{1-\frac{v^2}{c^2}}\;\Delta t}\,. \tag{A.18}$$

$\Delta\tau < \Delta t$ shows that moving clocks run more slowly, a phenomenon known as **time dilation.** More specifically, to the observer, a traveler’s clocks run more slowly. But, to the traveler, the observer’s clocks also run more slowly. This is because there is no preferred inertial frame.

This leads to the twin paradox where a twin travels and returns, but both can’t be younger. The resolution in Euclidean space is that the traveling twin must turn around, involving an acceleration, and that portion is not covered by equation (A.18).

**Length contraction** (or **Lorentz contraction**) occurs similarly. Consider a rod positioned on the $x'$-axis of frame K′ that has **proper length** (or **rest length**)

$$\ell_0 = x'_2 - x'_1\,, \tag{A4-1}$$

where $x'_1$ and $x'_2$ are the coordinates of the rod’s endpoints in frame K′. From equations (A.12),

$$x'_1 = \gamma\,(x_1 - v\,t)$$
$$x'_2 = \gamma\,(x_2 - v\,t)$$
$$\ell_0 = x'_2 - x'_1 = \gamma\,(x_2 - x_1)$$

Let $\ell = x_2 - x_1$. Then

$$\ell_0 = \gamma\,\ell = \frac{\ell}{\sqrt{1-\frac{v^2}{c^2}}}\,,\text{ or}$$

$$\boxed{\ell = \ell_0\sqrt{1-\frac{v^2}{c^2}}}\,: \tag{A.19}$$

Since $\ell < \ell_0$, a stationary observer sees that the rod is contracted, and a traveler moving along with the rod likewise sees a stationary rod as contracted.

There is no contraction in the *y* and *z* directions. Thus the volume V of a moving object is related to its rest volume $V_0$ by

$$\boxed{V = V_0 \sqrt{1 - \frac{v^2}{c^2}}} .$$

This fact must be taken into account when considering densities.

## A.5 Spacetime diagrams

Spacetime diagrams are either 2D or 3D spacetime graphs with one or two space dimensions suppressed. It is conventional to have the $t$ axis point vertically upward, and for the straight-line path of photons to be inclined at 45°. As explained for Figure A.1b, his is equivalent to using units in which $c = 1$ or using the coordinates $x^\mu$ defined by equations (A.3) where $x^0 = c\,t$.

Figure A.3 shows the **null cone at event O**. If event O is taken as the origin of an inertial reference system, then **the equation of the null cone is**

$$x^2 + y^2 + z^2 = c^2\, t^2 \,. \tag{A.20}$$

It shows the **future of O** and its **past**. It shows vectors localized at O as **timelike vectors**, **spacelike vectors**, and **null vectors** (or **lightlike vectors**):

$$\lambda^\mu \text{ is} \quad \begin{cases} \text{Timelike} \\ \text{Null or Lightlike} \\ \text{Spacelike} \end{cases} \quad \text{if } \eta_{\mu\nu}\,\lambda^\mu\,\lambda^\nu \begin{cases} > 0 \\ = 0 \\ < 0 \end{cases} \tag{A.21}$$

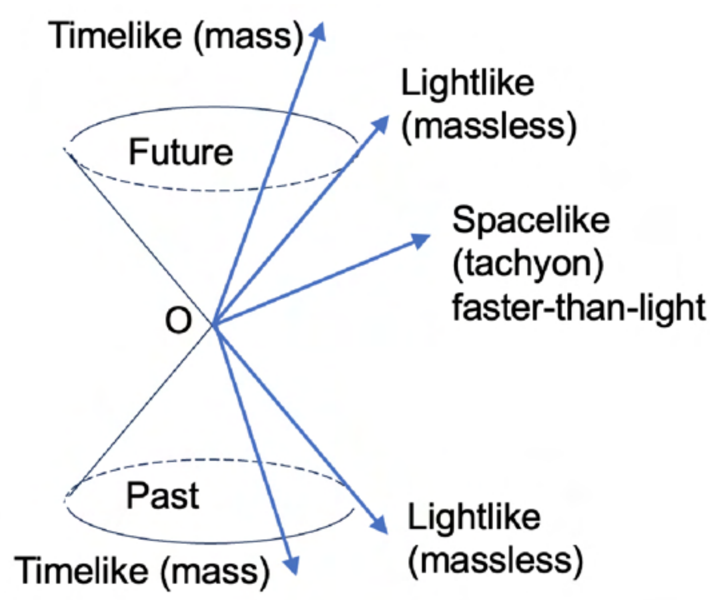



Figure A.3. Spacetime Light Cone

Timelike vectors are either **future-pointing** or **past-pointing**. Spacelike vectors are neither in the future nor the past.

Relativistic mechanics limits a particle with mass to speeds below *c*. Its worldline lies within the null cone. Its proper time

$$\tau \overset{(A.6)}{=} t\,\sqrt{1-\frac{v^2}{c^2}}$$

may be used to parameterize the word line:

$$x^\mu = x^\mu(\tau).$$

Its tangent vector

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$$

is called the **world velocity** of the particle. To justify calling $u^\mu$ a vector, we must show that definition (A1-3) is satisfied, that $u^{\mu'} = \Lambda^{\mu'}_{\nu}\, u^\nu$:

$$u^{\mu'} = \frac{\mathrm{d}x^{\mu'}}{\mathrm{d}\tau} \overset{(A1-1)}{=} \frac{\mathrm{d}\left(\Lambda^{\mu'}_{\nu}\, x^\nu\right)}{\mathrm{d}\tau} = \Lambda^{\mu'}_{\nu}\, \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = \Lambda^{\mu'}_{\nu}\, u^\nu$$

because $\Lambda^{\mu'}_{\nu}$ is a constant (with formula given in equation A1-20). ✓

Next,

$$\eta_{\mu\nu}\, u^\mu\, u^\nu = c^2 :$$

$$c^2\,(\mathrm{d}\tau)^2 \overset{(A.5)}{=} \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = \eta_{\mu\nu}\, u^\mu\, u^\nu\,(\mathrm{d}\tau)^2 \quad ✓$$

Since $\eta_{\mu\nu}\, u^\mu\, u^\nu = c^2 > 0$, by definition (A.21) $u^\mu$ is timelike, has magnitude *c*, and lies within the null cone at each event on the worldline. The tangent vector at each event on the world line of a photon is null. These situations are shown in Figure A.3.

Spacetime diagrams can be used to illustrate Lorentz transformations. For example, if we reverse the *x* and *t* axes of Figure A.1.b we get the boost spacetime diagram on LHS of Figure A.4 below. The dotted line represents the limit when $v = c = 1$. It has slope $c = 1/c = 1$. The RHS illustrates an event P. The horizontal dashed line represents constant *t*, so it represents simultaneous events in K. Similarly the slanted dashed line parallel to the $x'$-axis represents simultaneous events in K′. Event Q occurs after event P in frame K but it occurs before event P in frame K′. This could lead to some philosophical difficulties if one observer believes event P causes event Q.

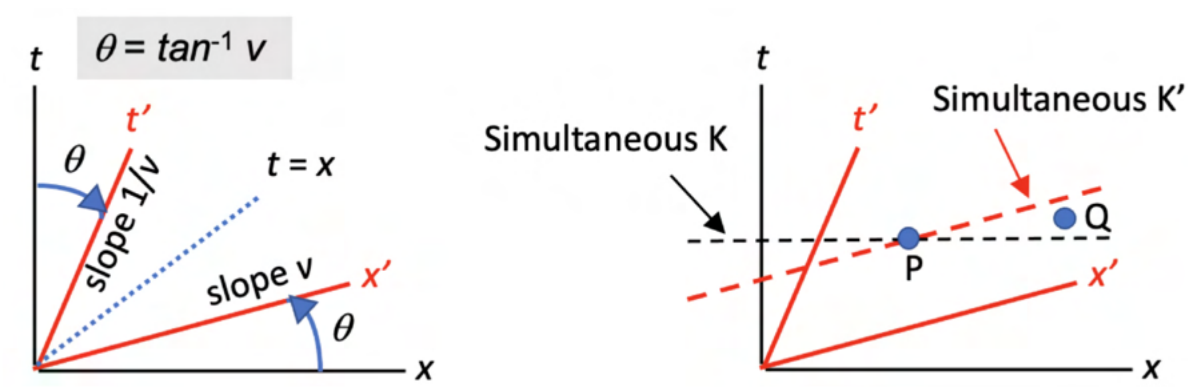



Figure A.4  Spacetime diagram of a boost

We can also determine the slopes of the $t'$ and $x'$ axes directly from the boost equations (A.12). The $x'$-axis represents $t' = 0$, which from equation (A.12) gives $t - \frac{v}{c}$ x = 0. Since we have labeled the time axis by *t* rather than *ct*, we have effectively set $c = 1$. So, $t = vx$ is the equation of the $x'$-axis, and it has slope *v*. Similarly, we can set $x' = 0$ in equation (A.12) to get $x - vt = 0$, or $t = \frac{1}{v}$ x as the equation of the $t'$-axis, and its slope is $\frac{1}{v}$.

## A.6 Some standard 4-vectors

Notation Space vectors are 3-vectors and spacetime vectors are 4-vectors. We denote

$$\lambda^\mu = (\lambda^0, \lambda^1, \lambda^2, \lambda^3) = (\lambda^0, \boldsymbol{\lambda}) \tag{A.22}$$

where the spatial part will be written in boldface.

We have already introduced the **coordinate vector** $x^\mu = (ct, x, y, z) = (ct, \boldsymbol{x})$. $x^\mu$ is a vector by equation (A1-1):

$$x^{\mu'} \overset{(A1-1)}{=} \Lambda^{\mu'}_{\nu}\, x^\nu. \qquad \checkmark$$

Contrast this with the Euclidean coordinate object $x^i = (x, y, z) = \boldsymbol{x}$, which is not a 3-vector, demonstrated in Example 1.4.2.

We have also already introduced the **4-velocity**, called the **world velocity** in section A.5, where we showed that it is a vector:

$$\boxed{u^\mu \equiv \frac{dx^\mu}{d\tau}}. \tag{A6-1}$$

The **coordinate velocity** object is

$$\boxed{v^\mu \equiv \frac{dx^\mu}{dt} = (c, \boldsymbol{v})}, \tag{A.23}$$

where $\boldsymbol{v}$ is the particle's 3-velocity. The same approach that showed $u^\mu$ is a vector proves that $v^\mu$ is *not* a vector. We compare them, one above the other, with $u^\mu$ first:

$$u^{\mu'} = \frac{dx^{\mu'}}{d\tau} \overset{(A1-1)}{=} \frac{d\left(\Lambda^{\mu'}_{\nu}\, x^\nu\right)}{d\tau} = \Lambda^{\mu'}_{\nu} \frac{dx^\nu}{d\tau} = \Lambda^{\mu'}_{\nu}\, u^\nu$$

$$v^{\mu'} = \frac{dx^{\mu'}}{dt'} \overset{(A1-1)}{=} \frac{d\left(\Lambda^{\mu'}_{\nu}\, x^\nu\right)}{d\,t'} = \Lambda^{\mu'}_{\nu} \frac{dx^\nu}{dt'} \neq \Lambda^{\mu'}_{\nu} \frac{dx^\nu}{dt} = \Lambda^{\mu'}_{\nu}\, v^\nu \qquad \checkmark$$

It may help in understanding this to observe that $\tau$ is Lorentz invariant, the same in all coordinate systems, but $t$ is not; it becomes $t'$ when the coordinate system changes.

Similar to equation (A.23), we have that

$$v^{\mu\prime} = \frac{dx^{\mu\prime}}{dt'} = (c\frac{dt'}{dt'}, \frac{d\boldsymbol{x}'}{dt'}) = (c, \boldsymbol{v}').$$

Also, observe that

$$\boxed{\frac{dt}{d\tau} = \gamma}: \tag{A6-2}$$

$$\frac{dt}{d\tau} \overset{(A.6)}{=} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \overset{(A.11)}{=} \gamma \qquad ✓$$

We can express $u^\mu$ in terms of $v^\mu$:

$$u^\mu = \frac{dx^\mu}{dt}\frac{dt}{d\tau} = \gamma\, v^\mu = \gamma\,(\text{c}, \boldsymbol{v})\,. \tag{A.24}$$

Also, take care that the partial and total derivatives with respect to $t$ are different. The partial derivatives, like $\frac{\partial z}{\partial x}$, relate to one coordinate with respect to another, and equal zero. In particular, $\frac{\partial x}{\partial t} = 0$. Partials occur when computing the Lagrangian and the Christoffel symbols $\Gamma_{\delta\nu\gamma}$. However, the total derivative relates to position and time. Thus, $\boldsymbol{v} = \frac{d\boldsymbol{x}}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right)$ is not zero unless a particle is stationary.

If we define

$$\gamma' \equiv \frac{dt'}{d\tau},$$

then, like unprimed equation (A.24), we have

$$u^{\mu\prime} = \gamma' v^{\mu\prime} = \gamma'(c, \boldsymbol{v}'):$$

$$u^{\mu\prime} = \frac{dx^{\mu\prime}}{d\tau} = (c\frac{dt'}{d\tau}, \frac{d\boldsymbol{x}'}{d\tau}) = (c\frac{dt'}{d\tau}, \frac{d\boldsymbol{x}'}{dt'}\frac{dt'}{d\tau}) = \gamma'(c, \boldsymbol{v}') = \gamma' v^{\mu\prime} \qquad ✓$$

Additionally, since $v' = -v$, we also get that

$$\gamma' = \frac{dt'}{d\tau} \overset{(A.6)}{=} \frac{1}{\sqrt{1 - \frac{v'^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \overset{(A.6)}{=} \frac{dt}{d\tau} = \gamma.$$

A **particle's 4-momentum** is defined as $\boxed{p^\mu \equiv m\, u^\mu}$ (A.25)

where $m$ is the particle's rest mass. It is a vector because it is the product of a scalar and a vector. (Compare to equation A.24 that does not make $v^\mu$ a vector because $\gamma$ is a variable, not a scalar).

The zeroth component of momentum of a particle is defined as

$$p^0 \equiv E\,/\,c, \tag{A6-3}$$

where $E$ is the particle's energy. So, we can express

$$p^\mu = (E/c,\ \boldsymbol{p}) \tag{A6-4}$$

where $\boldsymbol{p} \overset{(A.22)}{=} (p^1,\ p^2,\ p^3)$. (A.26)

Moving on from *particles*, the *object*

$$\boxed{k^\mu \equiv \left(\frac{2\,\pi}{\lambda},\ \boldsymbol{k}\right)} \tag{A.27}$$

is defined to be the **wave 4-vector of a photon**, where **$\boldsymbol{\lambda}$ is the wavelength** and

$$\boldsymbol{k} = \frac{2\,\pi}{\lambda}\,\boldsymbol{n}, \tag{A6-5}$$

where **$\boldsymbol{n}$ is a unit 3-vector in the direction of propagation**. If we let $\boldsymbol{w}$ be the velocity 3-vector of the photon, then $\boldsymbol{w} = c\,\boldsymbol{n}$ and we have that

$$k^\mu = (k^0,\ \boldsymbol{k}) = \frac{2\,\pi}{\lambda}\ (1,\ \frac{\boldsymbol{w}}{c})\ . \tag{A6-6}$$

We defer proof that $k^\mu$ is a *vector* until after we develop the Doppler shift formula in Section A.7.

Next, $\boldsymbol{k} \bullet \boldsymbol{k} = \dfrac{4\,\pi^2}{\lambda^2}$, and so we get

$$k^\mu\, k_\mu \overset{(1.26)}{=} \eta_{\mu\nu}\, k^\mu\, k^\nu = \eta_{00}\,\left(k^0\right)^2 + \boldsymbol{k} \bullet \boldsymbol{k}\ = -\,\frac{4\,\pi^2}{\lambda^2} + \frac{4\,\pi^2}{\lambda^2} = 0. \tag{A6-7}$$

That is, $k^\mu$ is null (i.e., lightlike), hence tangential to the photon's world line.

The **photon's 4-momentum** is defined as

$$\boxed{p^\mu \equiv \hbar\, k^\mu}\ , \tag{A.28}$$

where $\hbar = \frac{h}{2\pi}$ is Planck's reduced constant. $p^\mu = (p^0, \boldsymbol{p})$, where $\boldsymbol{p} = \hbar\, \boldsymbol{k}$, and $p^\mu$ is a vector because it is the product of a scalar and a vector.

Because $\kappa^\mu$ is null, a photon's 4-momentum is null:

$$\boxed{p^\mu\, p_\mu = 0}\ : \tag{A6-8}$$

$$p_\mu \overset{\text{(Th A0-1)}}{=} g_{\mu\nu}\, p^\nu = \hbar\, g_{\mu\nu}\, k^\nu = \hbar\, k_\mu \quad \Rightarrow \quad p^\mu\, p_\mu = \hbar^2\, k^\mu\, k_\mu \overset{\text{(A6-7)}}{=} 0\ . \quad ✓$$

Also,

$$p^0 = \frac{h}{\lambda}: \tag{A6-9}$$

$$p^0 \overset{\text{(A.28)}}{=} \hbar\, k^0 \overset{\text{(A.27)}}{=} \hbar\, \frac{2\pi}{\lambda} = \frac{h}{2\pi}\,\frac{2\pi}{\lambda} = \frac{h}{\lambda}\ . \quad ✓$$

In addition, just as we did in definition (A6-3) for a particle, we define the **zeroth component of photon momentum** as

$$p^0 \equiv \frac{E}{c}. \tag{A6-10}$$

We define the **photon's frequency** $\nu$ as:

$$\nu \equiv \frac{c}{\lambda}\ . \tag{A6-11}$$

Then, as in Quantum Mechanics,

$$\mathrm{E} = h\,\nu: \tag{A6-12}$$

$$\mathrm{E} \overset{\text{(A6-10)}}{=} c\, p^0 \overset{\text{(A6-9)}}{=} h\,\frac{c}{\lambda} \overset{\text{(A6-11)}}{=} h\,\nu\ .$$

This is not yet insightful because rest mass equation (A6-10) and 4-momentum equation (A.28) have been defined so as to ensure this. The energy formula from quantum mechanics is the motivation for defining $p^0 = \frac{E}{c}$ for photons. Since photon speed is the limiting case of particle speed, this also serves as motivation for having defined $p^0 = \frac{E}{c}$ in equation (A.6-3) for particles. Additional motivation will be given shortly, as well.

We can now express photon momentum as

$$p^\mu = (\frac{E}{c}, \boldsymbol{p}) \tag{A6-13}$$

where $\boldsymbol{p} = \hbar\, \boldsymbol{k}$, identical to the momentum equation (A6-3) for particles.

Returning to particles, in relativistic mechanics, **Newton's second law** is modified to

$$\mathrm{d}p^\mu / \mathrm{d}\tau = f^\mu \tag{A.29}$$

where $f^\mu$ is the **4-force** on the particle. This reflects that mass can change over time.

Letting $\boldsymbol{F}$ be the **3-force**, we define

$$\boxed{f^\mu \equiv \left(f^0,\ \gamma \boldsymbol{F}\right)}\ . \tag{A.30}$$

$f^\mu$ is a vector:

$$f^{\mu\prime} = \frac{\mathrm{d}p^{\mu\prime}}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}(\Lambda^{\mu\prime}_{\ \nu}\, p^\nu) \overset{(A5-1)}{=} \Lambda^{\mu\prime}_{\ \nu} \frac{\mathrm{d}p^\nu}{\mathrm{d}\tau} = \Lambda^{\mu\prime}_{\ \nu}\, f^\nu \quad ✓$$

Exercise A.6.2 $f^0 = \frac{\gamma}{c}\, \boldsymbol{F} \bullet \boldsymbol{v}$ (A.6.14)

Solution. $c^2\, \mathrm{d}\tau^2 \overset{(A.5)}{=} \eta_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu$ (A6-15)

$$\Rightarrow \quad c^2 \overset{(A6-15)}{=} \eta_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \overset{(A6-1)}{=} \eta_{\mu\nu}\, u^\mu\, u^\nu$$

$$\Rightarrow \quad 0 = \frac{\mathrm{d}c^2}{\mathrm{d}\tau} = \eta_{\mu\nu} \left(\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} u^\nu + u^\mu \frac{\mathrm{d}u^\nu}{\mathrm{d}\tau}\right) = \eta_{\nu\mu} \frac{\mathrm{d}u^\nu}{\mathrm{d}\tau} u^\mu + \eta_{\mu\nu} u^\mu \frac{\mathrm{d}u^\nu}{\mathrm{d}\tau} = 2\, \eta_{\mu\nu} u^\mu \frac{\mathrm{d}u^\nu}{\mathrm{d}\tau}$$

$= 2\, u_\mu \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}$ (since, by Example A.0.1 part (b), $\eta_{\mu\nu}$ lowers superscripts)

$$\Rightarrow \quad u_\mu\, f^\mu \overset{(A.29)}{=} u_\mu \frac{\mathrm{d}p^\mu}{\mathrm{d}\tau} \overset{(A.25)}{=} m u_\mu \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = 0\ .$$

Since $f^\mu \overset{(A.30)}{=} (f^0, \gamma \boldsymbol{F})$, then

$$0 = u_\mu\, f^\mu = u_0\, f^0 - \boldsymbol{u} \bullet \gamma \boldsymbol{F} \overset{(A.24)}{=} \gamma c\, f^0 - \gamma^2\, \boldsymbol{v} \bullet \boldsymbol{F}$$

$$\Rightarrow \quad f^0 = \frac{\gamma}{c}\, \boldsymbol{F} \bullet \boldsymbol{v} \quad ✓$$

■

Example A.6.1 Let K′ be an instantaneous rest frame. Then

$$p^{\mu\prime} = (mc, \mathbf{0}), \quad p^{0\prime} = mc, \quad \text{and} \quad \mathbf{p}' = \mathbf{0}. \tag{A6-16}$$

Solution. $\mathbf{v}' = 0$ and $\gamma \overset{(A.11)}{=} 1$

$$\Rightarrow \quad (p^{0\prime}, \mathbf{p}') = p^{\mu\prime} \overset{(A.25)}{=} m\, u^{\mu\prime} \overset{(A.24)}{=} m\,(\gamma c, \gamma \mathbf{v}') = (mc, \mathbf{0})$$ ■

From Example A.0.1,

$$p_{\mu'} = (p^{0\prime}, -\mathbf{p}') \overset{(A6-16)}{=} (mc, \mathbf{0}) \quad \text{and} \quad p_\mu = (p^0, -\mathbf{p}) \overset{(A6-3)}{=} \left(\frac{E}{c}, -\mathbf{p}\right). \tag{A6-17}$$

Since inner products are invariant,

$$p^\mu\, p_\mu = p^{\mu\prime}\, p_{\mu'}\,. \tag{A.31}$$

Since, by definition, $p^2 = \mathbf{p} \bullet \mathbf{p}$, we have

$$m^2 c^2 = (mc, \mathbf{0}) \bullet (mc, \mathbf{0}) \overset{(A6-16,\, A6-17)}{=} p^{\mu\prime}\, p_{\mu'} \overset{(A.31)}{=} p^\mu\, p_\mu$$

$$\overset{(A6-4,\, A6-17)}{=} \frac{E^2}{c^2} - \mathbf{p} \bullet \mathbf{p} = \frac{E^2}{c^2} - p^2$$

$\Rightarrow E^2 = m^2 c^4 + c^2 p^2$, or

$$\boxed{E = \sqrt{p^2\, c^2 + m^2\, c^4}}\,. \tag{A.32}$$

This is the formula that connects energy of a particle to its momentum and rest mass. When the particle is at rest, this becomes $E = mc^2$. The $p^2 c^2$ term is not found in Newtonian mechanics. ■

At non-relativistic speeds, the momentum equation reduces to the Newtonian formula for total energy in terms of potential and kinetic energy:

$$\left(\frac{E}{c}, \mathbf{p}\right) \overset{(A6-10)}{=} (p^0, \mathbf{p}) = p^\mu \overset{(A.25)}{=} m u^\mu \overset{(A.24)}{=} (\gamma m c, \gamma m \mathbf{v})\,.$$

$$\Rightarrow \quad \mathbf{p} = \gamma m \mathbf{v} \tag{A.33}$$

$$\Rightarrow \quad E = \gamma m c^2 = m c^2 \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = m c^2 + \frac{1}{2} m v^2 + \cdots\,. \tag{A.34}$$

where $v^2 = \boldsymbol{v} \cdot \boldsymbol{v}$. Equation (A.33) shows that $\boldsymbol{p} \approx m\boldsymbol{v}$ when $\boldsymbol{v}$ is small. Equation (A.34) shows that $E \approx mc^2 + \frac{1}{2}mv^2$ when $\boldsymbol{v}$ is small. That is, $E$ is the sum of the **rest energy** ($mc^2$) and the **kinetic energy** ($\frac{1}{2}mv^2$). (Equation (A.34) is a consequence of the power series expansion $(1-x)^{-a} = 1 + ax + \frac{1}{2}\,a\,(a\text{-}1)\,x^2 + \frac{1}{3!}\,a\,(a\text{-}1)\,(a\text{-}2)\,x^3 + ...$).

E $\approx mc^2$ is motivation for having *defined* $p^0 = \frac{E}{c}$ for particles. Additional motivation comes the desire to preserve the law of conservation of momentum in Relativity.

From consideration of simple collision problems in different frames, it turns out that we must define $\boldsymbol{p} = \gamma m\boldsymbol{v}$ rather than as $m\boldsymbol{v}$. We examine $p^0$ in that light.

$$\frac{dp^\mu}{d\tau} \overset{(A.29)}{=} f^\mu \overset{(A.30)}{=} \gamma\left(\boldsymbol{F} \cdot \frac{\boldsymbol{v}}{c}, \boldsymbol{F}\right) \;\Rightarrow\; \frac{dp^0}{d\tau} = \frac{\gamma}{c}\,\boldsymbol{F} \cdot \boldsymbol{v}\,.$$

$\boldsymbol{F} \cdot \boldsymbol{v}$ is the Newtonian rate of imparting energy of a 3-force into a particle. The corresponding relativistic rate would be $\gamma\,\boldsymbol{F} \cdot \boldsymbol{v}$. So, defining $E = c\,p^0$ yields

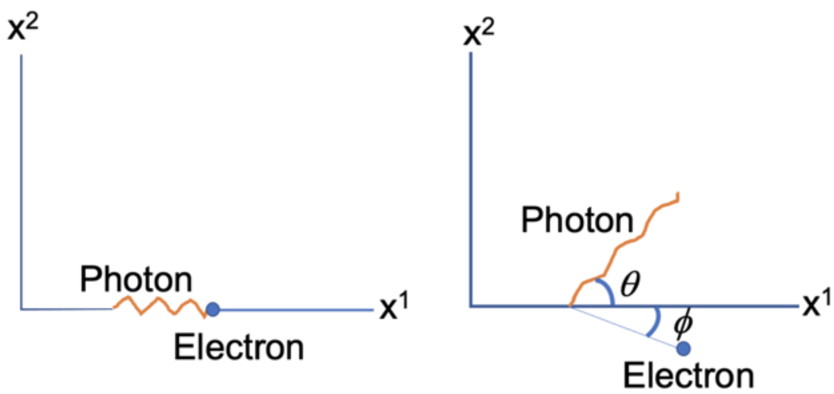
$$\frac{dE}{d\tau} = c\,\frac{dp^0}{d\tau} = \gamma\,\boldsymbol{F} \cdot \boldsymbol{v} \quad ✓$$

As a consequence, the single equation $p^\mu = \left(\frac{E}{c}, \boldsymbol{p}\right) = \text{constant}$ represents both conservation of energy (first term) and conservation of momentum (second term). This extends to a system of interacting particles having no external forces:

$$\sum_{\text{all particles}} p^\mu = \text{constant} \qquad \text{(A.35)}$$

This next example shows the computations for a simple particle collision problem.

Example A.6.2 In the **Compton effect**, a photon collides with a stationary electron. Assume the photon initially travels along the positive $x^1$-axis and that after the collision both particles travel in the $x^1x^2$-plane, making angles $\theta$ and $\phi$, respectively, with the the $x^1$-axis. Find the electron's velocity and the photon's frequency after the collision.




Solution. For the photon after-collision frequency, $\overline{\nu}$, we seek a formula dependent only on the pre-collision parameters along with the photon's after-collision deflection angle, $\theta$. It should not include the electron's post-collision speed $v$ or deflection angle, $\phi$.

Similarly for $v$, we seek a formula independent of $\theta$ and $\overline{\nu}$.

For the photon, $E \overset{(A6-12)}{=} h\,\nu$ and $p^\mu\, p_\mu \overset{(A6-8)}{=} 0$.

Before the collision:

$$p^\mu_{\text{ph}} \overset{(A6-3)}{=} \left(\frac{E}{c}, \boldsymbol{p}_{\text{ph}}\right) \overset{(A.28)}{=} \left(\frac{E}{c}, \hbar\, \boldsymbol{k}\right) \overset{(A6-5)}{=} \left(\frac{h\nu}{c}, \frac{h}{2\pi}\frac{2\pi}{\lambda}, 0, 0\right) \overset{(A6-12)}{=} \left(\frac{h\nu}{c}, \frac{h\nu}{c}, 0, 0\right)$$

$$p^\mu_{\text{el}} \overset{(A6-3)}{=} \left(\frac{E}{c}, \boldsymbol{p}_{\text{el}}\right) \overset{(A.32)}{=} \left(\frac{mc^2}{c}, 0, 0, 0\right) = (mc, 0, 0, 0)$$

After the collision:

$\overline{p}^\mu_{\text{ph}} = \left(\frac{h\overline{\nu}}{c}, \frac{h\overline{\nu}}{c}\cos\theta, \frac{h\overline{\nu}}{c}\sin\theta, 0\right)$ where $\overline{\nu}$ is photon's frequency after the collision

$\overline{p}^\mu_{\text{el}} \overset{(A.25)}{=} m\, u^\mu_{\text{el}} \overset{(A.24)}{=} (m\gamma c, m\gamma v\cos\phi, -m\gamma v\sin\phi, 0)$ where $u^\mu_{\text{el}}$ is the electron's world velocity after collision, and $v$ is the electron's speed after collision

We can express the conservation of system momentum by

$p^\mu_{\text{ph}} + p^\mu_{\text{el}} = \overline{p}^\mu_{\text{ph}} + \overline{p}^\mu_{\text{el}}$, or

(1) $\frac{h\nu}{c} + mc = \frac{h\overline{\nu}}{c} + \gamma mc$

(2) $\frac{h\nu}{c} = \frac{h\overline{\nu}}{c}\cos\theta + \gamma mv\cos\phi \iff \cos\phi = \frac{h}{\gamma mvc}(\nu - \overline{\nu}\cos\theta)$

(3) $0 = \frac{h\overline{\nu}}{c}\sin\theta - \gamma mv\sin\phi \iff \sin\phi = \frac{h\overline{\nu}\sin\theta}{\gamma mvc}$

When we eliminate $\phi$ and $v$, the result (see Exercise A.6.3) is

$$\boxed{\bar{\nu} = \frac{\nu}{1 + \frac{h\nu}{mc^2}\,(1 - cos\theta)}}, \tag{A.36}$$

which is the **Compton scattering formula**, and is the solution for the photon's frequency.

As for the electron's velocity, eliminating $\theta$ and $\bar{\nu}$ yields a 4th degree equation in $v$ that has a solution in theory but that I can't solve (see my Exercise A.6.3). ■

## A.7 Doppler effect

The text book never proves that the wave vector $\kappa$ is a vector, but nonetheless uses that fact to develop the Doppler formula. I do the reverse. I develop the Doppler formula (for a general 3-space boost, not for just an $x$-boost), and then I use the Doppler formula to prove that $\kappa$ is a vector. I adapt various methods of Frobenius, cited earlier.

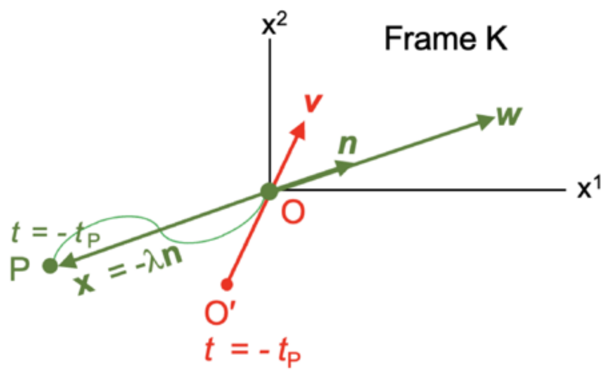



Figure A.7.1

In the figure above, the $x^3$- axis has been suppressed to reduce clutter. Figure A.7.1shows inertial frame K having spatial origin O, and it shows the spatial origin O′, of inertial frame K′, moving at 3-space velocity $\boldsymbol{v}$ with respect to K. We set time so that the spatial origins O and O′ coincide at $t = t' = 0$.

The Doppler Shift formula represents the ratio of the wavelength seen by an observer at the origin of frame K to the wavelength seen by the emitter at point P traveling along with frame K′. The ratio is not affected if the origin O′ is shifted by an offset so that it does not pass through point O. Thus, it is sufficient to derive the ratio using this general homogeneous boost rather than a more general one having non-zero offset.

Suppose a source of radiation is traveling with frame K′. A photon is emitted by the source at point P, which is one wavelength $\lambda$ from O. The photon is detected by an observer at O when the photon arrives, at time 0. We denote the photon velocity vector by $\boldsymbol{w}$. Frame K is the observer frame, and frame K′ is the emitter frame.

We set $t_P$ to be the photon time of flight from P to O. Then $\lambda = c\, t_P$, and the 4-vector in frame K for the photon emission event is $x^\mu = (x^0, \boldsymbol{x}) = (-c\, t_P, -\lambda\, \boldsymbol{n}) = (-\lambda, -\frac{\lambda}{c} \boldsymbol{w})$, where

$\boldsymbol{x}$ is the Frame K coordinates of P and $\boldsymbol{n}$ is the unit vector in the direction of photon motion. In order to reach O at $t = 0$, the photon would have been emitted at time $t = -t_P$.

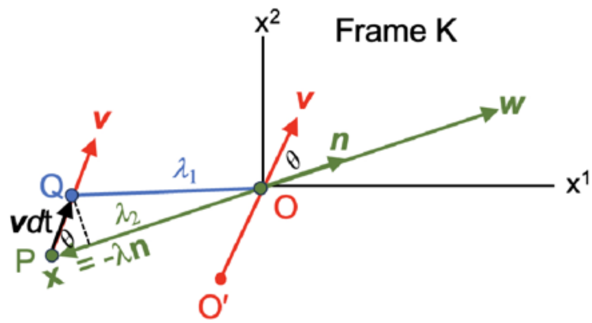



Figure A.7.2

Figure A.7.2 shows the radiation source moving with frame K′ from point P to point Q, and then emitting a second photon at time $t = -t_P + \mathrm{d}t$ . The spatial coordinates of Q are, thus, $\boldsymbol{x} + \boldsymbol{v}\,\mathrm{d}t$. We let $\lambda_1$ denote the distance from Q to O, and we set $\lambda_2 \equiv \lambda - v\,\mathrm{d}t\cos\theta$, as shown, where $\theta$ is the acute angle between $\boldsymbol{v}$ and $\boldsymbol{w}$ .

Let $t = \hat{t}_1$ be the arrival time at O of the second photon. Since the first photon arrives at $t = 0$, the difference in arrival times of the photons is

$$\mathrm{d}\hat{t} \equiv \hat{t}_1 - 0 = \hat{t}_1 = -t_P + \mathrm{d}t + \frac{\lambda_1}{c} = \mathrm{d}t + \frac{\lambda_1 - c\,t_P}{c} = \mathrm{d}t + \frac{\lambda_1 - \lambda}{c}.$$

W have that $\boldsymbol{v} \bullet \boldsymbol{w} = v\,c\cos\theta$, and from Figure A.7.2 we see that

$$\lambda_1 \approx \lambda_2 = \lambda - v\,\mathrm{d}t\cos\theta.$$

$$\Rightarrow \quad \mathrm{d}\hat{t} \approx \mathrm{d}t - \frac{v}{c}\,\mathrm{d}t\cos\theta = \left(1 - \frac{v}{c}\cos\theta\right)\mathrm{d}t \overset{(A1-13)}{=} \left(1 - \frac{v}{c}\cos\theta\right)\gamma\,\mathrm{d}t' = \left(1 - \frac{\boldsymbol{v} \bullet \boldsymbol{w}}{c^2}\right)\gamma\,\mathrm{d}t'$$

Since $\lambda = c\,t$, wavelength is proportional to time. Therefore,

$$\frac{\mathrm{d}\hat{t}}{\mathrm{d}t'} = \frac{\mathrm{d}t\,(\text{observed})}{\mathrm{d}t\,(\text{emitted})} = \frac{\lambda\,(\text{observed})}{\lambda\,(\text{emitted})} = \frac{\lambda}{\lambda'}.$$

This yields the Doppler Shift formula:

$$\boxed{\frac{\lambda}{\lambda'} = \gamma\left(1 - \frac{\boldsymbol{v} \bullet \boldsymbol{w}}{c^2}\right)} \qquad \text{(A7-1)}$$

If the source P is approaching the observer, then $\theta = 0$ and

$$\boxed{\frac{\lambda}{\lambda'} = \gamma\left(1 - \frac{v}{c}\right) = \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}}} \tag{A.40}$$

The observed wavelength is blue-shifted.

If the source is receding from the observer, then $\theta = \pi$ and

$$\boxed{\frac{\lambda}{\lambda'} = \gamma\left(1 + \frac{v}{c}\right) = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}} \tag{A.41}$$

The observed wavelength is red-shifted.

Both of these Doppler effects have their counterpart in classical physics. However, the next case is new, only in Relativity. The angle $\theta$ can be $\pm\frac{\pi}{2}$. For example, if the source travels perpendicularly to the *x*, *y*, or *z*-axis and intersects the axis, then $\theta = \pm\frac{\pi}{2}$ as the source crosses the axis.

When this happens, we get

$$\boxed{\frac{\lambda}{\lambda'} = \gamma = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}}}, \tag{A.42}$$

a redshift called the **transverse Doppler effect**.

The last topic for this section is to use the Doppler formula to prove that the wave vector $k^\mu$ is indeed a vector, as promised in the prior section . Recall equation (A6-6):

$$k^\mu = (k^0, \boldsymbol{k}) = \frac{2\pi}{\lambda}(1, \frac{\boldsymbol{w}}{c}),$$

where $\boldsymbol{w}$ is the velocity 3-vector of the photon, as shown in Figure A.7.2, above. We have to show that if K′ is an inertial frame moving at velocity $\boldsymbol{v}$ with respect to inertial frame K, then $k^\mu$ satisfies the Lorentz equation

$$\kappa^{\mu'} = \Lambda^{\mu'}_{\ \nu}\,\kappa^{\nu} \ \text{ where } \ \Lambda^{\mu'}_{\ \nu} \overset{\text{(A1–20)}}{=} \begin{pmatrix} \gamma & -\frac{\gamma\,\boldsymbol{v}^{\mathrm{T}}}{c} \\ -\frac{\gamma\,\boldsymbol{v}}{c} & \mathrm{I} + \frac{\gamma^2}{c^2\,(\gamma+1)}\,\mathbf{v}\mathbf{v}^{\mathrm{T}} \end{pmatrix}.$$

However, like the Doppler shift, the wave frequency is not affected by an offset. So, it is sufficient to use the homogeneous Lorentz transformation that corresponds to Figure A.7.2 and show that

$$\kappa^{\mu'} = \tilde{\Lambda}^{\mu'}_{\ \nu}\,\kappa^{\nu} \ \text{ where } \ \tilde{\Lambda}^{\mu'}_{\ \nu} \overset{\text{(A1–17)}}{=} \gamma \begin{pmatrix} 1 & \frac{-\boldsymbol{v}^{\mathrm{T}}}{c} \\ \frac{-\boldsymbol{v}}{c} & \mathrm{I} \end{pmatrix}. \qquad \text{(A7.2)}$$

Thus, we must show

$$\frac{2\,\pi}{\lambda'}\begin{pmatrix} 1 \\ \frac{\boldsymbol{w}'}{c} \end{pmatrix} = \kappa^{\mu'} = \tilde{\Lambda}^{\mu'}_{\ \nu}\,\kappa^{\nu} = \gamma \begin{pmatrix} 1 & \frac{-\boldsymbol{v}^{\mathrm{T}}}{c} \\ \frac{-\boldsymbol{v}}{c} & \mathrm{I} \end{pmatrix} \frac{2\,\pi}{\lambda}\begin{pmatrix} 1 \\ \frac{\boldsymbol{w}}{c} \end{pmatrix} = \frac{2\,\pi}{\lambda}\,\gamma \begin{pmatrix} 1 - \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{c^2} \\ \frac{-\boldsymbol{v} + \boldsymbol{w}}{c} \end{pmatrix} \qquad \text{(A7.3)}$$

This results in two equations that we must show:

$$\frac{2\,\pi}{\lambda'} = \frac{2\,\pi}{\lambda}\,\gamma\,(1 - \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{c^2}) \qquad \text{(A7-4)}$$

and

$$\frac{2\,\pi}{\lambda'}\,\frac{\boldsymbol{w}'}{c} = \frac{2\,\pi}{\lambda}\,\gamma\,\frac{-\boldsymbol{v} + \boldsymbol{w}}{c} \qquad \text{(A7-5)}$$

Equation (A7-4) is simply a restatement of the Doppler shift formula (A7-1):

$$\frac{2\,\pi}{\lambda'} \overset{\text{(A7–1)}}{=} \frac{2\,\pi}{\lambda}\,\gamma\,(1 - \frac{\boldsymbol{v} \cdot \boldsymbol{w}}{c^2}) \qquad \checkmark$$

We can generate equation (A7-5) from the homogeneous Lorentz *coordinate transformation* (A1-15) for frame K′ having constant velocity $\boldsymbol{v}$ :

$$t' = \gamma\,(t - \frac{\boldsymbol{v} \cdot \boldsymbol{x}}{c^2})$$
$$\boldsymbol{x}' = \gamma\,(\boldsymbol{x} - \boldsymbol{v}\,t) \qquad \text{(A1-22)}$$

Taking derivatives of both equations yields

$$\mathrm{d}t' = \gamma\,(\mathrm{d}t - \frac{\boldsymbol{v}}{c^2} \cdot \mathrm{d}\boldsymbol{x}) \quad \text{and} \quad \mathrm{d}\boldsymbol{x}' = \gamma\,(-\,\boldsymbol{v}\,\mathrm{d}t + \mathrm{d}\boldsymbol{x})\,. \qquad \text{(A7-6)}$$

Expressing $\boldsymbol{w}'$ as $\frac{\mathrm{d}\boldsymbol{x}'}{\mathrm{d}t'}$, we compute $\boldsymbol{w}'$ by dividing the two (A7-6) equations:

$$\boldsymbol{w}' = \frac{\mathrm{d}\boldsymbol{x}'}{\mathrm{d}t'} = \frac{-\boldsymbol{v}\,\mathrm{d}t + \mathrm{d}\boldsymbol{x}}{\mathrm{d}t - \frac{\boldsymbol{v}}{c^2} \bullet \mathrm{d}\boldsymbol{x}} = \frac{-\boldsymbol{v} + \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}}{1 - \frac{\boldsymbol{v}}{c^2} \bullet \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}} = \frac{-\boldsymbol{v} + \boldsymbol{w}}{1 - \frac{\boldsymbol{v} \bullet \boldsymbol{w}}{c^2}} . \qquad \text{(A7-7)}$$

This yields desired equation (A7-5),

$$[\frac{2\pi}{\lambda'}]\,[\frac{\boldsymbol{w}'}{c}] \overset{\text{(A7–4, A7–7)}}{=} [\frac{2\pi}{\lambda}\gamma\,(1 - \frac{\boldsymbol{v} \bullet \boldsymbol{w}}{c^2})]\,[\frac{-\boldsymbol{v} + \boldsymbol{w}}{\left(1 - \frac{\boldsymbol{v} \bullet \boldsymbol{w}}{c^2}\right)}\,\frac{1}{c}] = \frac{2\pi}{\lambda}\gamma\,\frac{-\boldsymbol{v} + \boldsymbol{w}}{c} \qquad \checkmark$$

and completes the proof that $k^\mu$ is a vector. ■

Note 1. The Lorentz coordinate transformation equations (A1-22) applies to any *point*, but when we examine $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$, what *velocity* is represented? The answer is that we represented $\boldsymbol{w}$ above, but in fact we could have had any velocity in mind. For example, had we set $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$ to represent $\boldsymbol{v}$, we would have derived $\boldsymbol{v}' = \frac{2\pi}{\lambda}\gamma\,\frac{-\boldsymbol{v} + \boldsymbol{v}}{c} = 0$ which, indeed, is the velocity of a fixed point in the K′ frame with respect to the K′ frame.

Note 2. Equations (A1-22) were derived using $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v}$, which is proper because frame K′ moves with velocity $\boldsymbol{v}$ with respect to frame K. Above, we switched to $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{w}$. This is permissible because once the equations for *position* are established, then the derivative equations (A7.6) are free to represent any velocity for that point, not just $\boldsymbol{v}$. As further justification, the usual derivation of equations (A1-22) in other text books does not use $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$. It is computed from rotations and boosts based on $\boldsymbol{v}$, not $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$.

## A.8 Electromagnetism

Let

$\boldsymbol{E}$ = electric field intensity
$\boldsymbol{B}$ = magnetic induction
$\rho$ = charge density (i.e., charge per unit volume)
$\rho_0$ = proper charge density (i.e., charge per unit *rest* volume)
$\boldsymbol{J}$ = current density
$\mu_0$ = permeability of free space
$\varepsilon_0$ = permittivity of free space

where

$$\mu_0\, \varepsilon_0 = 1 / c^2 \,. \tag{A.47}$$

Maxwell's equations in free space in differential equation form are

$$\nabla \bullet \boldsymbol{B} = 0 \tag{A.43}$$

$$\nabla \bullet \boldsymbol{E} = \rho / \varepsilon_0 \tag{A.44}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{A.45}$$

$$\nabla \times \boldsymbol{B} = \mu_0 \boldsymbol{J} + \mu_0 \varepsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} \tag{A.46}$$

where

$$\nabla = \frac{\partial}{\partial x}\boldsymbol{i} + \frac{\partial}{\partial y}\boldsymbol{j} + \frac{\partial}{\partial z}\boldsymbol{k} \tag{A8-1}$$

$$\nabla \bullet \boldsymbol{B} = \frac{\partial B^x}{\partial x} + \frac{\partial B^y}{\partial y} + \frac{\partial B^z}{\partial z} \tag{A8-2}$$

$$\nabla \times \boldsymbol{B} = \left(\frac{\partial B^z}{\partial y} - \frac{\partial B^y}{\partial z}\right)\boldsymbol{i} + \left(\frac{\partial B^x}{\partial z} - \frac{\partial B^z}{\partial x}\right)\boldsymbol{j} + \left(\frac{\partial B^y}{\partial x} - \frac{\partial x}{\partial y}\right)\boldsymbol{k} \tag{A8-3}$$

and similarly for $\boldsymbol{E}$.

We can also express curl **B** $\equiv \nabla \times$**B** = $\begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ B^x & B^y & B^z \end{vmatrix}$

and, so,

$$(\text{curl } \nabla)\phi = (\nabla\times\nabla)\phi = \nabla\times(\nabla\phi) = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\phi}{\partial x} & \frac{\partial\phi}{\partial y} & \frac{\partial\phi}{\partial z} \end{vmatrix} = 0 \qquad \text{(A8-4)}$$

for any twice differentiable potential $\phi$.

The vector fields $\boldsymbol{B}$ and $\boldsymbol{E}$ can be expressed in terms of a vector potential $\boldsymbol{A}$ and a scalar potential $\varphi$ :

$$\boldsymbol{B} = \nabla\times\boldsymbol{A} \quad \text{and} \quad \boldsymbol{E} = -\nabla\varphi - \frac{\partial\boldsymbol{A}}{\partial t}. \qquad \text{(A8-5)}$$

The book states that equations (A.43) and (A.45) are then satisfied:

$$\nabla\bullet\boldsymbol{B} = \nabla\bullet(\nabla\times\boldsymbol{A}) = \frac{\partial}{\partial x}\left(\frac{\partial A^z}{\partial y} - \frac{\partial A^y}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial A^x}{\partial z} - \frac{\partial A^z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial A^y}{\partial x} - \frac{\partial A^x}{\partial y}\right) = 0 \ ✓$$

$$\nabla\times\boldsymbol{E} = -(\nabla\times\nabla)\varphi - \nabla\times\frac{\partial\boldsymbol{A}}{\partial t} \overset{\text{(A8-4)}}{=} 0 - \frac{\partial}{\partial t}(\nabla\times\boldsymbol{A}) \overset{\text{(A8-5)}}{=} -\frac{\partial}{\partial t}(\boldsymbol{B}) \quad ✓$$

The potentials are not uniquely determined by Maxwell's equations: $\boldsymbol{A}$ can be replaced by $\boldsymbol{A} + \nabla\varphi$ and $\varphi$ by $\varphi - \frac{\partial\psi}{\partial t}$ where $\psi$ is arbitrary. Such transformations are known as **gauge transformations** and allow one to choose $\boldsymbol{A}$ and $\varphi$ so that the following condition, the **Lorentz gauge condition**, is satisfied:

$$\nabla\bullet\boldsymbol{A} + \mu_0\varepsilon_0\frac{\partial\varphi}{\partial t} = 0 \qquad \text{(A.49)}$$

First, we show that equations (A.43) and (A.45) are still satisfied.

$$\boldsymbol{B} \overset{\text{(A8-5)}}{=} \nabla\times(\boldsymbol{A} + \nabla\varphi)$$

$$\nabla\bullet\boldsymbol{B} = \nabla\bullet[\nabla\times(\boldsymbol{A} + \nabla\varphi)] = \nabla\bullet(\nabla\times\boldsymbol{A}) + \nabla\bullet(\nabla\times\nabla\varphi)$$

$$\overset{\text{(A8-4)}}{=} \nabla\bullet(\nabla\times\boldsymbol{A}) + 0 = \frac{\partial}{\partial x}\left(\frac{\partial A^z}{\partial y} - \frac{\partial A^y}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial A^x}{\partial z} - \frac{\partial A^z}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial A^y}{\partial x} - \frac{\partial A^x}{\partial y}\right)$$

$$= 0 \ ✓$$

$$\boldsymbol{E} \overset{\text{(A8-5)}}{=} -\nabla\left(\varphi - \frac{\partial\psi}{\partial t}\right) - \frac{\partial}{\partial t}(\boldsymbol{A} + \nabla\phi)$$

$$\nabla\times\boldsymbol{E} = -(\nabla\times\nabla)\left(\varphi - \frac{\partial\psi}{\partial t}\right) - \nabla\times\frac{\partial(\boldsymbol{A} + \nabla\varphi)}{\partial t}$$

$$\overset{\text{(A8-4)}}{=} 0 - \frac{\partial}{\partial t}(\nabla\times\boldsymbol{A}) - \frac{\partial}{\partial t}[(\nabla\times\nabla)\,\phi)]$$

$$\overset{\text{(A8-5, A8-4)}}{=} -\frac{\partial}{\partial t}(\boldsymbol{B}) - 0 = -\frac{\partial\boldsymbol{B}}{\partial t} \quad ✓$$

Next we proceed to choose $\boldsymbol{A}$ and $\varphi$ so that equation (A.49) is satisfied. We wish to find $\boldsymbol{A}$ and $\phi$ such that $\nabla \bullet (\boldsymbol{A} + \nabla\phi) + \mu_0\varepsilon_0 \frac{\partial}{\partial t}(\varphi - \frac{\partial \psi}{\partial t}) = 0$; i.e., such that $\nabla\bullet\boldsymbol{A} + \nabla\bullet\nabla\phi = -\mu_0\varepsilon_0\left[\frac{\partial}{\partial t}\varphi - \frac{\partial^2\psi}{\partial t^2}\right]$. Presumably we should first seek $\phi$ such that $\nabla\bullet\nabla\phi = \mu_0\varepsilon_0\frac{\partial^2\psi}{\partial t^2}$ and then find $\boldsymbol{A}$ such that $\nabla\bullet\boldsymbol{A} = -\mu_0\varepsilon_0\frac{\partial\phi}{\partial t}$.
But, note that the LHS's involve space coordinates while RHS's involve just the time coordinate. (✓ ??)

The remainder of this section leads eventually to the development of Maxwell equations, and the observation that they are Lorentz invariant but not Galilean invariant and so represent 4-tensors without modification:

$\boldsymbol{E} = (E^1, E^2, E^3)$, $\boldsymbol{B} = (B^1, B^2, B^3)$, $j^\mu = (\rho\, c, \boldsymbol{J}) = \rho\, v^\mu = (\gamma\, \rho_0)\, v^\mu = \rho_0\, u^\mu$,

$$(F_{\mu\nu}) = \begin{pmatrix} 0 & -\frac{E^1}{c} & -\frac{E^2}{c} & -\frac{E^3}{c} \\ \frac{E^1}{c} & 0 & B^3 & -B^2 \\ \frac{E^2}{c} & -B^3 & 0 & B^1 \\ \frac{E^3}{c} & B^2 & -B^1 & 0 \end{pmatrix} \tag{A.54}$$

$$F^{\mu\nu}{}_{,\nu} = \mu_0\, j^\mu \tag{A.55}$$

$$F_{\mu\nu,\sigma} + F_{\nu\sigma,\mu} + F_{\sigma\mu,\nu} = 0 \tag{A.56}$$

Contrast this with Newtonian mechanics equations that are Galilean invariant and so require modification to be generalized to 4-tensors in Special Relativity.

# Chapter 2 (part 2) Spacetime and Gravitation

## 2.5 The spacetime of general relativity

Definition As discussed in Appendix A, the **spacetime of special relativity** is a 4-dimensional pseudo-Riemannian manifold with a global coordinate system in which the metric tensor can be expressed as

$$(\eta_{\mu\nu}) \equiv \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{2.5-1}$$

A spacetime coordinate system that satisfies (2.5-1) is called a **Cartesian coordinate system**. In a Cartesian coordinate system,

$$x^0 = c\,t, \quad x^1 = x, \quad x^2 = y, \quad \text{and} \quad x^3 = z.$$

Recall that at any point P in general relativity spacetime there is a local coordinate system in which $(\Gamma^{\mu}_{\nu\,\sigma})_{\,\mathrm{P}} = 0$ and $(x_\mu)_{\,\mathrm{P}} = (0,0,0,0)$. This implies that $(\partial_\sigma\, g_{\mu\nu})_{\,\mathrm{P}} = 0$, which implies that the spacetime metric tensor $g_{\mu\nu}$ is approximately equal to the Cartesian metric tensor $\eta_{\mu\nu}$ to the first order. That is, $g_{\mu\nu} \approx \eta_{\mu\nu}$:

Using multivariate Taylor series expansion,

$$\begin{aligned} g_{\mu\nu} &= \eta_{\mu\nu} + (\partial_\sigma\, g_{\mu\nu})_{\mathrm{P}}\, (x^\sigma - 0) + \tfrac{1}{2}\, (\partial_\rho\, \partial_\sigma\, g_{\mu\nu})_{\mathrm{P}}\, (x^\rho - 0)\, (x^\sigma - 0) + \ldots \\ &= \eta_{\mu\nu} + \tfrac{1}{2}\, (\partial_\rho\, \partial_\sigma\, g_{\mu\nu})_{\mathrm{P}}\, x^\rho x^\sigma + \ldots \\ &\approx \eta_{\mu\nu} \text{ to the first order} \qquad \checkmark \end{aligned}$$

The **spacetime of general relativity** is a 4-dimensional pseudo-Riemannian manifold that has a system of local Cartesian coordinates in which $g_{\mu\nu} \approx \eta_{\mu\nu}$ to the first order at each point P:

$$g_{\mu\nu} \approx \eta_{\mu\nu} + \tfrac{1}{2}\, (\partial_\rho\, \partial_\sigma\, g_{\mu\nu})_{\mathrm{P}}\, x^\rho x^\sigma \qquad (\mu, \nu, \rho, \sigma = 0, 1, 2, 3) \tag{2.66}$$

When $g_{\mu\nu} = \eta_{\mu\nu}$ we say the (local) **spacetime is flat**. Otherwise it is **curved**.

Observe that $\eta_{\mu\nu}$ is not a tensor in curved general relativity spacetime but only in flat special relativity spacetime. For example, if $\Lambda^{\rho'}_{\mu}$ represents a transformation (at some point P) from Cartesian to spherical coordinates, then $\eta_{\mu\nu} \neq \Lambda^{\rho'}_{\mu} \Lambda^{\sigma'}_{\nu}\, \eta_{\rho'\,\sigma'}$ because, as shown in Section 1.2, the space part of $(\eta_{\mu'\nu'})$ is not the identity matrix:

$$\Lambda_i^{k'}\Lambda_j^{\ell'}\eta_{k'\ell'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -r^2 & 0 \\ 0 & 0 & -r^2 \sin^2\theta \end{pmatrix} \neq \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = (\eta_{i'j'})$$

So, $\eta_{\mu\nu}\tau^\nu$ is not a tensor and $\eta$ cannot be used to raise and lower indices. Raising and lowering indices is still done with $g$:

$$\tau_\mu = g_{\mu\nu}\tau^\nu, \quad \tau^\mu = g^{\mu\nu}\,\tau_\nu\,, \quad \text{and} \quad g_{\mu\sigma}\,g^{\sigma\nu} = \delta^\nu_\mu.$$

Because general relativity spacetime is locally approximately like special relativity spacetime, **special relativity results can be generalized to general relativity using the following principles**:

1. If a physical quantity can be defined as a Cartesian tensor in special relativity, then we can define it in exactly the same way in a local Cartesian coordinate system.
2. Any valid tensor equation in special relativity can be converted to a valid tensor equation in general relativity simply by replacing
    a. Partial differentiation (denoted by a comma) by covariant differentiation (denoted by a semicolon)
    b. Total derivatives (d/d$u$) along a curve by absolute derivatives (D/d$u$)
    c. $\eta_{\mu\nu}$ by $g_{\mu\nu}$ (i.e., use the true metric tensor rather than an approximation)

Example 1  Maxwell's equations (A.55 and A.56) in special relativity were expressed

$$F^{\mu\nu}{}_{,\nu} = \mu_0\, j^\mu$$

$$F_{\mu\nu,\sigma} + F_{\nu\sigma,\mu} + F_{\sigma\mu,\nu} = 0 \tag{2.67}$$

In general relativity this becomes

$$F^{\mu\nu}{}_{;\nu} = \mu_0\, j^\mu$$

$$F_{\mu\nu\,;\sigma} + F_{\nu\sigma\,;\mu} + F_{\sigma\mu\,;\nu} = 0 \tag{2.68}$$

Example 2  In special relativity the world velocity of a particle having rest mass $m > 0$ is

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \tag{2.5-2}$$

where proper time $\tau$ for the particle is defined by

$$c^2\,(\mathrm{d}\tau)^2 \overset{(A.5)}{=} \eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu.$$

Its equation of motion is defined in terms of the 4-force $f^\mu$ by

$$\frac{\mathrm{d}\,p^\mu}{\mathrm{d}\tau} \overset{(A.29)}{=} f^\mu$$

where

$$p^\mu \overset{(A.25)}{=} m u^\mu\,.$$

Implementation in general relativity spacetime is still

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \quad \text{and} \quad p^\mu = m u^\mu \quad \text{(because we showed these to be 4-vectors)}$$

but now **proper time is defined by the general relativity line element**

$$\boxed{c^2\,(\mathrm{d}\tau)^2 \equiv g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu} \tag{2.69}$$

and the equation of motion becomes

$$\frac{\mathrm{D}\,p^\mu}{\mathrm{d}\tau} = f^\mu. \tag{2.70}$$

In the case of a free particle (i.e., $f^\mu = 0$), this reduces to

$$\boxed{\frac{\mathrm{d}^2\,x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\nu\sigma}\,\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\,\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} = 0} \tag{2.71}$$

$$0 \overset{(2.70)}{=} \frac{\mathrm{D}\,p^\mu}{\mathrm{d}\tau} = m\,\frac{\mathrm{D}\,u^\mu}{\mathrm{d}\tau} \overset{(2.45)}{=} m\left(\frac{\mathrm{d}^2\,x^\mu}{\mathrm{d}\tau^2} + \Gamma^\mu_{\nu\sigma}\,\frac{\mathrm{d}x^\sigma}{\mathrm{d}\tau}\,\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\right) \quad ✓ \quad \text{(Exercise 2.5.2)}$$

The path of a free particle in flat spacetime is a straight line. By definition (2.12) this generalizes to the path of a free particle in curved spacetime following a geodesic having the proper time of a particle as its affine parameter. Equations (2.71) are the **equations of motion for a free particle (with mass) in curved spacetime.**

For a photon (or any massless particle), there is no change in proper time $\tau$ along its path, so $\tau$ cannot be used as in equation (2.71) to parameterize its worldine. Since in special relativity a photon travels in a straight line, then in curved space it travels along an affinely parameterized geodesic. That is, there is an affine parameter $u$ that parameterizes the worldine of a photon:

$$\frac{\mathrm{d}^2\,x^\mu}{\mathrm{d}u^2} + \Gamma^\mu_{\nu\sigma}\,\frac{\mathrm{d}x^\sigma}{\mathrm{d}u}\,\frac{\mathrm{d}x^\nu}{\mathrm{d}u} \overset{(2.12)}{=} 0. \tag{2.72}$$

Note. We do not actually solve for $u$. Rather, we simply use its existence to write the photon equation of motion (2.72).

The photon also satisfies the next equation, which is equivalent to its speed being $c$:

$$g_{\mu\nu}\,\frac{dx^\mu}{du}\,\frac{dx^\nu}{du} = 0: \tag{2.73}$$

$$v = c \;\Leftrightarrow\; c^2 = v^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 \;\Leftrightarrow\; c^2\,dt^2 = dx^2 + dy^2 + dz^2$$

$$\Leftrightarrow\; g_{\mu\nu}\,dx^\mu\,dx^\nu \overset{(2.69)}{=} c^2\,d\tau^2 \overset{(A.2)}{=} c^2\,dt^2 - dx^2 + dy^2 + dz^2 = 0 \quad \checkmark$$

The next topic, characterization of vectors $\lambda^\mu$ in a manifold, has already been done:

$$\begin{pmatrix} \textbf{timelike} \\ \textbf{null},\ \text{or}\ \textbf{lightlike} \\ \textbf{spacelike} \end{pmatrix} \text{if } g_{\mu\nu}\,\lambda^\mu\lambda^\nu \begin{pmatrix} >\ 0 \\ =\ 0 \\ <\ 0 \end{pmatrix}$$

However, it should be pointed out that in curved spacetime the null cone is also curved and so null vectors lie in the tangent spaces at the given points and not in the manifold.

At any point on the path of a particle with mass, its world velocity $u^\mu = \dfrac{dx^\mu}{d\tau}$ is a tangent vector to the path, and equation (2.69) tells us this tangent vector is timelike. So, a particle with mass follows a timelike path through spacetime, and a free particle follows a timelike geodesic. A photon follows a null geodesic path and equation (2.73) tells us that the tangent vectors to its path are null.

Note: By a “free” particle we mean one that is under no external force. In this context, gravity is not regraded as a force but as a curvature of spacetime that influences geodesics. If we regard gravity as a force, then a “free” particle is one that experiences only the gravitational force.

By comparing equation (2.71) with its special relativity counterpart $\dfrac{d^2 x^\mu}{d\tau^2} = 0$, we see that the connection coefficients play an important role in explaining gravitational effects. In this sense, the metric tensor field, from which the connection coefficients are defined, carry the gravitational context of spacetime. How these metric tensor fields are determined by the distribution of matter will be discussed in chapter 3.

## 2.6 Newtonian gravitation and fluid dynamics

In this section we derive several fundamental equations in Newtonian physics to which corresponding equations of General Relativity must reduce. The equations are

1. Poisson's equation (the field equation for Newtonian gravitation) (2.6-11)
2. Classical continuity equations for a perfect fluid (2.6-12)
3. Euler's classical equation of motion for a perfect fluid (2.6-15)

We begin with Newton's 3 laws of motion.

**First Law** A body at rest remains at rest, and a body in motion travels with a constant speed and direction unless disturbed by an external force.

In general relativity this becomes: a free particle follows a geodesic path through space-time. Using a coordinate system where $\Gamma^{\mu}_{\nu\sigma} \approx 0$, this reduces to $\frac{d^2 x^\mu}{d\tau^2} = 0$. For $v \ll c$, $\frac{d\tau}{dt} \approx 1$ so the geodesic equation yield the familiar $\frac{d^2 x^i}{dt^2} = 0$ ($i$ = 1,2,3) straight line motion.

**Second Law** $\boldsymbol{F} = \mathrm{m}\,\frac{d^2 \boldsymbol{x}}{dt^2}$ is rendered in special relativity to allow mass to change with speed: $\boldsymbol{F} = \frac{d\boldsymbol{p}}{dt}$. In general relativity, gravity is introduced via the connection coefficients: $\boldsymbol{f} \overset{(2.70)}{=} \frac{D\boldsymbol{p}}{d\tau}$.

**Third Law** To every action there is an equal and opposite reaction. In general relativity this is modified to include motion along a geodesic. General relativity ignores any curvature that might be caused by the particle in motion. That is, the particle is treated as a test particle that has no influence on the body producing the gravitational field. Thus, general relativity does not handle the case, for example, of massive binary stars that revolve around each other, though there have been attempts at approximation.

The following additional reviews of Newtonian energy, gravitational potential, and fluid flow are not in the book.

Consider a particle moving from point A to point B, and denote points in between as $x$.

Set

**$K$ = Kinetic energy at $x$**
**$U$ = Potential energy at $x$**
**$E$ = Total energy at $x$**
**$W$ = Work performed to move the particle from point A to point B**
**$V$ = Potential energy per unit mass at $x$**

Law of Conservation of Energy  Total energy is preserved in moving a particle from point A to point B.

Think of moving a particle directly upwards from its resting spot on the Earth. The $x$-direction is then "up" (or radially outward), so

$E = K + U$ and $\Delta E = 0$, or $\Delta U = - \Delta K$.

The work performed is the product of the magnitude of the force applied and the distance moved: $W = F\,\Delta x$. $F$ can represent the the conglomeration of many different types of forces, or it can represent just a single type of force like gravitational force or electromagnetic force.

The gravitational force is **conservative**, meaning that total work performed in moving an object from A to B depends only on the location of points A and B and not on the path between them. We restrict our attention to conservative forces. Let $m$ be the mass of the point being moved, $v$ be its velocity, and $a$ its acceleration. Let $x$ be the direction of movement from A to B. Then

$$a = \frac{dv}{dt} = \frac{\partial v}{\partial x}\frac{dx}{dt} + \frac{\partial v}{\partial y}\frac{dy}{dt} + \frac{\partial v}{\partial z}\frac{dz}{dt} = v\,\frac{dv}{dx},$$

$$\int_A^x F\,dx = \int_A^x ma\,dx = \int_A^x mv\,\frac{dv}{dx}\,dx = \int_{v_A}^v mv\,dv = \frac{1}{2}mv^2 - \frac{1}{2}mv_A^2$$

$$= \Delta K = -\Delta U = -[U(x) - U(A)]$$

Let $A$ be a reference point where we assign $U(A) = 0$. Then

$$\int_A^x F\,dx = -U(x) \quad \text{and} \quad F(x) = -\frac{d}{dx}U(x)\,. \tag{2.6-1}$$

Since $a = \frac{F}{m}$, then $\int_A^x a\,dx = -\frac{U(x)}{m}$. Define the potential energy per unit mass by

$$V(x) \equiv \frac{U(x)}{m}.$$

Then

$$\boxed{\int_A^x a\,dx \;=\; -V \;\text{ and }\; a = -\frac{dV}{dx}}\,. \qquad (2.6\text{-}2)$$

Formula (2.6-2) tells us that for an acceleration field *a*, there is a *potential function V* such that $a = -\dfrac{dV}{dx}$. Compare this to equation (2.6–-1) that tells us that for a force field *F* there is a *potential function U* such that $F = -\dfrac{dU}{dx}$. If the field is the gravitational field, then ***U*** is the **gravitational potential** and ***V*** is the **gravitational potential per unit mass** or the **gravitational acceleration potential**.

Like a force field, *a* can represent acceleration due to an aggregation of many different types of forces. If we ignore all forces except gravity, then we can write $g = -\dfrac{dV}{dx}$.

In 3 dimensions, equation (2.6-2) generalizes to

$$\boldsymbol{a}(x,y,z) = -\left(\frac{\partial V}{dx}\boldsymbol{i} + \frac{\partial V}{dy}\boldsymbol{j} + \frac{\partial V}{dz}\boldsymbol{k}\right) = -\nabla V,$$

which can be expressed either in terms of components or as a vector:

$$\boxed{\frac{d^2 x^i}{dt^2} = -\partial_i V} \qquad \text{or} \qquad \boxed{\boldsymbol{a} = -\nabla V} \qquad (2.6\text{-}3)$$

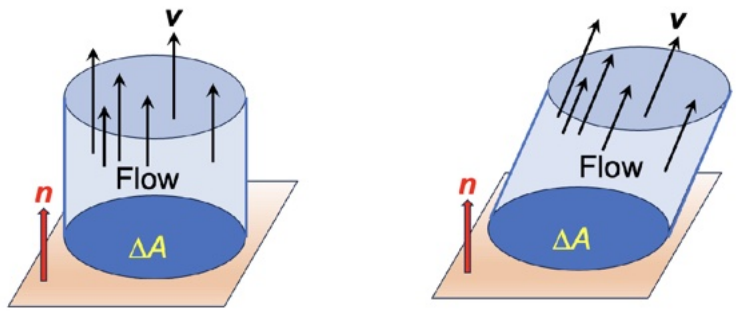
Equation (2.6-3) is the **Newtonian equation of motion for a particle moving in a gravitational field of potential *V***.

Next we develop equations of motion for fluids, which include liquids, gases, and plasma (a mixture of electrons and isotopes, produced, for example, by the sun).

For a fluid, **flux** Φ is defined as the rate of fluid that flows through a surface; i.e., the amount of mass that flows during a unit of time. We start by finding flux for a small part of a surface and then integrate to generate flux for the full surface.

We can consider a small surface, $\delta$S, to be flat and the velocity of the fluid to be the same at every point of the surface. Flux can be defined in terms of either volume or mass. Let the fluid have density $\rho$ and speed *v.* Let $\Delta A$ be the area of $\delta$S and $\hat{\boldsymbol{n}}$ the unit normal vector to $\delta$S that has the same orientation as the flow, as shown below.

The left figure shows the case where the velocity of the fluid is in the direction $\hat{\boldsymbol{n}}$.




$V_\Phi = v \times$ (1 unit of time) $\times\, \Delta A$ is the volume of fluid that flows in a unit of time
$\Delta\Phi = \rho V_\Phi = \rho v\, \Delta A$ is the mass of the fluid that flows through $\delta$S in a unit of time

If the direction of flow is not the same as $\hat{\boldsymbol{n}}$, we get the skewed cylinder on the right with base $\Delta A$ and height $v\, cos\theta$, where $\theta$ is the angle between the velocity $\boldsymbol{v}$ and $\hat{\boldsymbol{n}}$.

$V_\Phi = v\, \Delta A\, cos\theta = \Delta A\; \boldsymbol{v} \bullet \hat{\boldsymbol{n}}$

$\Delta\Phi = \rho\, \Delta A\; \boldsymbol{v} \bullet \hat{\boldsymbol{n}}$

If we define a vector $\boldsymbol{\delta A}$ whose magnitude is the area $\Delta A$ and whose direction is normal to $\delta$S, then $\boldsymbol{\delta A} = \Delta A\; \hat{\boldsymbol{n}}$. So, we can rewrite $\Delta\Phi$ as

$\Delta\Phi = \rho\; \boldsymbol{v} \bullet (\Delta A\; \hat{\boldsymbol{n}}) = \rho \boldsymbol{v} \bullet \boldsymbol{\delta A}$

Just as breaking the full surface S into small, flat surfaces $\delta$S, adding them up, and taking the limit gives the surface area of S,

$A = \iint_S \mathrm{d}\boldsymbol{A}$,

adding up the small fluxes, and taking the limit gives

$\Phi = \iint_S \rho \boldsymbol{v} \bullet \mathrm{d}\boldsymbol{A}$.

If we define the vector field as $\boldsymbol{F}$, this can be written

$$\boxed{\Phi = \iint_S \boldsymbol{F} \bullet \mathrm{d}\boldsymbol{A}}\;. \qquad (2.6\text{-}4)$$

Any integral of this form is called a **flux integral**. (This $\boldsymbol{F}$ has units of momentum but we can consider the concept of flux when $\boldsymbol{F}$ has other units, like force, below.)

Gravitation is related to flux and potential. Assume the mass $M$ that generates the gravitational force is a point mass located at the origin. Newton's formula for the force extended on a small particle of mass $m$ located at a distance $r$ from the origin is

$$F = \frac{G M m}{r^2}\ , \tag{2.6-5}$$

where $G$ is **Newton's universal gravitational constant**. Equation (2.6-5) is known as **Newton's equation for universal gravitation**. Note that $G M / r^2$ must have units of acceleration. The **gravitational potential energy** $U$ is the work required to move the particle from some reference distance to a given distance, $r$. The reference point for gravitational potential energy is set to ∞, and the reference value is set to zero. That is, $U(\infty) \equiv 0$. Since positive work is required to move the particle from $r$ to ∞, we have that, for all $x$, $U(x) < 0$. We express $U$ as

$$U = -\frac{G M m}{r}:$$

$$U \overset{(2.6\text{–}1)}{=} \int_{\infty}^{r} F\, \mathrm{d}r = \int_{\infty}^{r} \frac{GMm}{r^2} \mathrm{d}r = -GMm\left(\frac{1}{r} - 0\right) = -\frac{G M m}{r}.$$

The force on a point mass located on a sphere of radius $r$ about the central mass $M$ is expressed as

$$\boldsymbol{a}(\boldsymbol{r}) \equiv -\frac{F}{m}\ \boldsymbol{e}_r = -\frac{G M}{r^2}\ \boldsymbol{e}_r\ ,$$

where $\boldsymbol{e}_r$ is a radial unit vector. The vector $\boldsymbol{a}$ is called the **gravitational field** or, sometimes, the **gravitational acceleration** (because $G M / r^2$ has units of acceleration). The “field” symbol $\boldsymbol{F}$ is often used instead of $\boldsymbol{a}$.

Let S be a solid ball of radius $r$ about the origin. Let its volume be denoted by $V_S$ and let ∂S be its surface. ∂S is a sphere of radius $r$ with surface area $A = 4\pi r^2$. Let $\boldsymbol{A}$ be a vector of magnitude $A$ that is normal to ∂S. Its unit direction vector is $\boldsymbol{e}_r$. By equation (2.6-4), the total flux of the gravitational field $\boldsymbol{F}$ across the surface ∂S is $\iint_{\partial S} \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{A}$ .

Claim $\iint_{\partial S} \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{A} = -4\pi GM$ : (2.6-6)

$$\iint_{\partial S} \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{A} = \iint_{\partial S} \left(-\frac{G M}{r^2}\ \boldsymbol{e}_r\right) \cdot (\boldsymbol{e}_r\ \mathrm{d}A) = -\frac{G M}{r^2} \iint_{\partial S} \boldsymbol{e}_r \cdot \boldsymbol{e}_r\ \mathrm{d}A = -\frac{G M}{r^2} \iint_{\partial S} \mathrm{d}A$$

$$= -\frac{G M}{r^2}\ 4\pi r^2 = -4\pi GM \quad ✓$$

By the Divergence Theorem, certain integrals over a surface can be expressed as integrals over its enclosed volume:

$$\iint_{\partial S} \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{A} = \iiint_S \nabla \cdot \boldsymbol{a} \, \mathrm{d}V_S \,. \qquad (2.6\text{-}7)$$

Let $\rho$ be the **mass density** (i.e., mass per unit volume) of a point of S. Then the gravitational mass $M$ can be expressed

$$M = \iiint_S \rho \, \mathrm{d}V_S \,. \qquad (2.6\text{-}8)$$

Claim $\nabla \cdot \boldsymbol{a} = -4\pi G\rho$ : (2.6-9)

$$\iiint_S \nabla \cdot \boldsymbol{a} \, \mathrm{d}V_S \overset{(2.6-7)}{=} \iint_{\partial S} \boldsymbol{a} \cdot \mathrm{d}\boldsymbol{A} \overset{(2.6-6)}{=} -4\pi GM \overset{(2.6-8)}{=} -4\pi G \iiint_S \rho \, \mathrm{d}V_S$$

$$= -\iiint_S 4\pi G\rho \, dV_S$$

$$\Rightarrow \quad \nabla \cdot \boldsymbol{a} = -4\pi G\rho \qquad ✓$$

Since gravity is a conservative force, by equation (2.6-2) there is a **gravitational acceleration potential** V such that

$$\boldsymbol{a} = -\nabla V \,. \qquad (2.6\text{-}10)$$

Thus, the gravitational acceleration potential, $V$, satisfies

$$\boxed{\nabla^2 V = 4\pi G\rho} \; : \qquad (2.6\text{-}11)$$

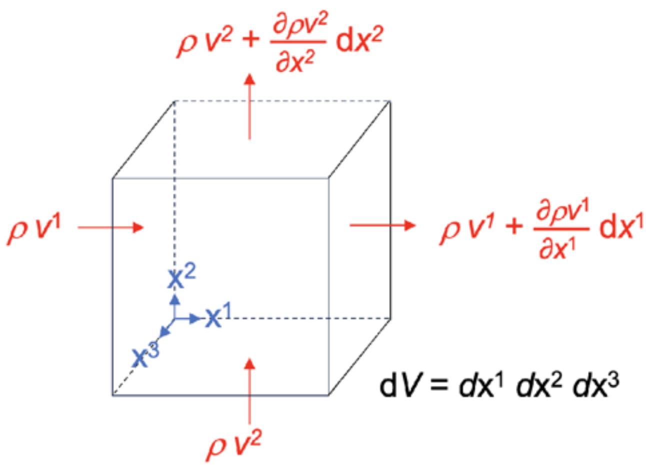
$$\nabla^2 V = \nabla \cdot \nabla V \overset{(2.6-10)}{=} -\nabla \cdot \boldsymbol{a} \overset{(2.6-9)}{=} 4\pi G\rho \qquad ✓$$

Equation (2.6-11) is known as **Poisson's equation**. It is the **field equation for Newtonian gravitation** to which Einstein's 16 field equations must reduce.

Next, we examine fluid flow to generate two more classical equations to which curved spacetime equations must reduce.

The most precise treatment of fluid flow, called **Lagrangian analysis**, tracks each individual particle over time. A simpler treatment, **Eulerian analysis**, focuses on the properties over time of each location ($x$,$y$,$z$) within a fluid, and tracks particles for only a short time interval d$t$. We begin with the continuity equation.

The continuity equation pertains to perfect fluids. It reflects the fact that mass is conserved in conventional Newtonian mechanics. The equation is developed by adding up the rate at which mass is flowing into and out of a control volume (depicted below as a cube) and setting the net mass in-flow equal to the net-outflow. In the figure below, $\rho$ is the (time-dependent) density of the fluid and $V_i$ is the volume of mass flowing in the *i*-direction.




First consider fluid flowing through the box with speed $v^1$ in the $x^1$-direction, from left to right. The area of the left face, and also the right face, is $dx^2 dx^3$. The amount of mass that flows into the box through the left face in a unit of time is

$$\rho\, v^1 dx^2 dx^3,$$

and the amount that flows out of the box through the right face is the flow in plus the change,

$$\rho\, v^1 dx^2 dx^3 + \frac{\partial(\rho\, v^1)}{\partial x^1} dx^1 (dx^2 dx^3).$$

The change in mass flow, in minus out, is

$$-\frac{\partial(\rho\, v^1)}{\partial x^1} dx^1 dx^2 dx^3.$$

Similar formulas apply to flow in the $x^2$ and $x^3$ directions. Keeping in mind that the volume of the box is $dx^1 dx^2 dx^3$, the total change is mass is

$$\frac{\partial \rho}{\partial t} dx^1 dx^2 dx^3 = \left[-\frac{\partial(\rho\, v^1)}{\partial x^1} - \frac{\partial(\rho\, v^2)}{\partial x^2} - \frac{\partial(\rho\, v^3)}{\partial x^3}\right] dx^1 dx^2 dx^3$$

$$\Rightarrow \quad \frac{\partial \rho}{\partial t} + \frac{\partial(\rho\, v^1)}{\partial x^1} + \frac{\partial(\rho\, v^2)}{\partial x^2} + \frac{\partial(\rho\, v^3)}{\partial x^3} = 0.$$

$$\Rightarrow \quad \frac{\partial \rho}{\partial t} + \left(\frac{\partial}{\partial x^1}\boldsymbol{i} + v^2 \frac{\partial}{\partial x^2}\boldsymbol{j} + v^3 \frac{\partial}{\partial x^2}\boldsymbol{k}\right) \bullet \rho\, (\, v^1 \boldsymbol{i} + v^2 \boldsymbol{j} + v^3 \boldsymbol{k}\, )$$

This is **the classical continuity equation for a perfect fluid**:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0. \qquad (2.6\text{-}12)$$

Euler's Equation is the other main equation in Newtonian mechanics for computing fluid flow. This equation describes flow in a perfect fluid. A fluid that can be completely characterized by its rest frame mass density $\rho$ and isotropic pressure $P$ (same in any direction) is called a **perfect fluid**. Real fluids are "sticky" and contain (and conduct) heat. In perfect fluids these possibilities are neglected. Specifically, perfect fluids have no shear stresses, viscosity, or heat conduction.

Perfect fluid flow is **steady**, meaning that velocity $\boldsymbol{v}$ at location $(x,y,z)$ is always the same. Of course, a *particle* can have different velocities during its flow, but once the particle reaches a location $(x,y,z)$, it will at that point have the same velocity as any other particle that later (or earlier) reaches this location.

The parameters describing Eulerian fluid motion at each location are its density $\rho$ and velocity $\boldsymbol{v}$. The cause of changes in fluid motion is force.

A force cannot be sustained by a single particle of a fluid, only by a surface. Furthermore, if the fluid is at rest and is to remain at rest, the force must be applied at a right angle to the surface (or particles will slide). Thus, the parameter for force in Euler's equation is pressure, $P$, which is defined as the magnitude of the normal force per unit area.

Derivation of Euler's equation starts with the same first step as the continuity equation except that pressure $P$ is substituted for mass $\rho A$. That is, the force due to constant pressure on a surface is simply the pressure times the area of the surface. So, the magnitude of the force on the left face in the figure above is

$$P\,\mathrm{d}x^2\mathrm{d}x^3,$$

and on the right face it is

$$\left[P + \frac{\partial P}{\partial x^1}\mathrm{d}x^1\right](\mathrm{d}x^2\mathrm{d}x^3).$$

Therefore, the net force in the $x$-direction is

$$-\frac{\partial P}{\partial x^1}\mathrm{d}x^1\mathrm{d}x^2\mathrm{d}x^3\boldsymbol{i},$$

where $\boldsymbol{i}$ is the unit vector pointing along the positive $x$-axis. The pressure per unit volume in the $x$-direction is

$$-\frac{\partial P}{\partial x^1}\boldsymbol{i},$$

the net force divided by the volume. Repeating this in the $y$ and $z$ directions yields a net force per unit volume

$$\boldsymbol{F} = -\left(\frac{\partial P}{\partial x^1}\boldsymbol{i} + \frac{\partial P}{\partial x^2}\boldsymbol{j} + \frac{\partial P}{\partial x^3}\boldsymbol{k}\right) = -\nabla P.$$

From Newton’s Second Law, we have force per unit volume = mass per unit volume times acceleration:

$$\boldsymbol{F} = \rho\,\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}.$$

Combining these yields

$$\rho\,\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\nabla P. \qquad (2.6\text{-}13)$$

The last step is to calculate the acceleration, $\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t}$, of the fluid particles. This is a bit tricky because a particle in the control volume can be subject to both internal and external forces. There could be acceleration $\frac{\partial \mathbf{v}}{\partial t}$ (note the partial derivative notation) as a particle moves within the inertial frame coordinate system. There could be further acceleration if the box itself is moving. For example, if the box is immersed in a circular flow, there will be centripetal acceleration.

To compute the acceleration of a particle, we need to find the rate of change of the velocity due to both internal and external forces:

$$\begin{aligned}
\frac{d\boldsymbol{v}}{dt} &= \lim_{\Delta t\to 0}\frac{\boldsymbol{v}\left(t+\Delta t,\, x+v^1\Delta t,\, y+v^2\Delta t,\, z+v^3\Delta t\right)}{\Delta t} = \frac{\partial \mathbf{v}}{\partial t} + v^1\frac{\partial \boldsymbol{v}}{\partial x^1} + v^2\frac{\partial \boldsymbol{v}}{\partial x^2} + v^3\frac{\partial \boldsymbol{v}}{\partial x^2} \\
&= \frac{\partial \mathbf{v}}{\partial t} + \left[\left(v^1\boldsymbol{i} + v^2\boldsymbol{j} + v^3\boldsymbol{k}\right)\bullet\left(\frac{\partial}{\partial x^1}\boldsymbol{i} + v^2\frac{\partial}{\partial x^2}\boldsymbol{j} + v^3\frac{\partial}{\partial x^2}\boldsymbol{k}\right)\right]\boldsymbol{v} \\
&= \frac{\partial \mathbf{v}}{\partial t} + (\boldsymbol{v}\bullet\nabla)\,\boldsymbol{v}. \qquad (2.6\text{-}14)
\end{aligned}$$

The first term reflects acceleration of the particle within the inertial frame and the second term reflects acceleration due to motion of the fluid body. The second term is non-linear and is the source of many difficulties in fluid mechanics.

Combining equations (2.6-13) and (2.6-14) yields **Euler's classical equation of motion for a perfect fluid:**

$$\rho \left[ \frac{\partial \mathbf{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\, \boldsymbol{v} \right] = -\nabla P \qquad (2.6\text{-}15)$$

## 2.7 Gravitational potential and the geodesic

A successful formulation of the 16 field equations of general relativity must reduce to to Poisson's equation, the field equation for Newtonian gravitation. What is a field equation? We defined fields for *spaces* in Section 1.6. We now extend the definition to manifolds.

Definitions  A **field** is an assignment of a scalar, vector, or tensor value to every point in manifold. A **field equation** is an equation that specifies the partial derivatives of the field at every point. Thus, integration of the partials along a curve between two points allows us to obtain the field value at the 2nd point from the 1st. We restrict our attention to conservative fields so that the integration result is independent of the the path between two points. Einstein's general relativity field equations are field equations that relate gravity to spacetime curvature.

As an example, a gravitational field requires specification of partials of 2nd order since gravity is an acceleration. We can imagine a field having complex hills and valleys that require higher dimensional partials in order to successfully obtain the value of nearby points by path integration. So, field equations are not tied to partials of 2nd order.

Poisson's equation, being the field equation for Euclidean gravity, simply requires 2nd order partials with respect to the three spatial axes. Recall Poisson's equation:

$$\nabla^2 V = 4\pi G\rho .$$

Einstein's gravitational field equations for spacetime also must include the second order partials, but partials with respect to time as well as space at every point ($x^\mu$). The first step in the process of deriving Einstein's field equations is to develop equation (2.83), the subject of the this section.

Suppose we have a coordinate system in which the metric tensor field is given by

$$g_{\mu\nu} \equiv \eta_{\mu\nu} + h_{\mu\nu} , \tag{2.74}$$

where

$$(\eta_{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} = (\eta^{\mu\nu}). \tag{2.7-1}$$

The aim of this section is to relate the $g_{00}$ component of the metric tensor (2.74) to the Newtonian gravitational potential when particle speed $\frac{dx^i}{dt}$ << $c$. In Section 2.5 we developed equation (2.71) for the motion for a free particle following a geodesic in curved spacetime with curvature specified by $g_{\mu\nu}$. In this section, we find the Euclidean 3-space equation of motion (2.7-13) for slowly moving particles to which this equation reduces, and we simplify the portion of the equations of motion that is due to the gravitational potential, leading to equation (2.81). Then, we generate equation (2.83), relating $g_{00}$ to the gravitational potential.

To approximate flat spacetime, the $h_{\mu\nu}$ in equation (2.74) are assumed to be small, 1st order terms with 2nd order terms that can be neglected:

$h_{\mu\nu} \approx 0$ to the 1st order. (2.7-2)

The contravariant version of equation (2.74) is then

$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$ to first order: (2.7-3)

Suppose $g^{\mu\nu} = \eta^{\mu\nu} + f^{\mu\nu}$ where $f^{\mu\nu}$ is a small (1st order) term.

Since $h_{\sigma\rho}$ is also of 1st order, the product $f^{\mu\nu} h_{\sigma\rho}$ is of 2nd order.

Since $g$ is invertible,

$\delta^\mu_\nu = g^{\mu\sigma} g_{\sigma\nu} = (\eta^{\mu\sigma} + f^{\mu\sigma})(\eta_{\sigma\nu} + h_{\sigma\nu}) = \delta^\mu_\nu + \eta^{\mu\sigma} h_{\sigma\nu} + f^{\mu\sigma}\eta_{\sigma\nu} + f^{\mu\sigma} h_{\sigma\nu}$

$\approx \delta^\mu_\nu + \eta^{\mu\sigma} h_{\sigma\nu} + f^{\mu\sigma}\eta_{\sigma\nu}$ to first order

$\Rightarrow \quad f^{\mu\sigma}\eta_{\sigma\nu} \approx -\eta^{\mu\sigma} h_{\sigma\nu}$

Because $f^{\mu\sigma} h_{\sigma\nu}$ and $h^{\mu\sigma} h_{\sigma\nu}$ are 2nd order terms, and $g^{\mu\sigma}$ raises subscripts,

$\Rightarrow \quad f^{\mu\rho} = f^{\mu\sigma}\delta^\rho_\sigma = f^{\mu\sigma} g^{\nu\rho} g_{\sigma\nu} = f^{\mu\sigma}(\eta_{\sigma\nu} + h_{\sigma\nu})\, g^{\nu\rho} = (f^{\mu\sigma}\eta_{\sigma\nu} + f^{\mu\sigma} h_{\sigma\nu})\, g^{\nu\rho}$

$\approx f^{\mu\sigma}\eta_{\sigma\nu}\, g^{\nu\rho} \approx -\eta^{\mu\sigma} h_{\sigma\nu}\, g^{\nu\rho} \approx -(\eta^{\mu\sigma} h_{\sigma\nu} + h^{\mu\sigma} h_{\sigma\nu})\, g^{\nu\rho}$

$= -(\eta^{\mu\sigma} + h^{\mu\sigma})\, h_{\sigma\nu}\, g^{\nu\rho} = -g^{\mu\sigma} g^{\nu\rho}\, h_{\sigma\nu} = -h^{\mu\rho}$

or $\quad f^{\mu\nu} \approx -h^{\mu\nu}$

$\Rightarrow \quad g^{\mu\nu} = \eta^{\mu\nu} + f^{\mu\nu} \approx \eta^{\mu\nu} - h^{\mu\nu}$ ✓

Claim:

$\Gamma^\mu_{\nu\sigma} \approx \frac{1}{2}\,\eta^{\mu\rho}(\partial_\nu h_{\rho\sigma} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma}) \approx 0$ to the 1st order: (2.78)

Since $\eta_{\mu\nu}$ is a constant, $\partial_\sigma \eta_{\mu\nu} = 0$, and we see that

$$\partial_\sigma g_{\mu\nu} \overset{(2.74)}{=} \partial_\sigma h_{\mu\nu}. \qquad (2.7\text{-}4)$$

$$\Gamma^\mu_{\nu\sigma} \overset{(2.13)}{=} g^{\mu\rho}\,\Gamma_{\rho\nu\sigma} \overset{(2.33)}{=} g^{\mu\rho}\,\frac{1}{2}\,(\partial_\nu g_{\rho\sigma} + \partial_\sigma g_{\nu\rho} - \partial_\rho g_{\nu\sigma})$$

$$\overset{(2.7-4)}{=} \frac{1}{2} g^{\mu\rho}(\partial_\nu h_{\rho\sigma} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma})$$
$$\overset{(2.7-3)}{\approx} \frac{1}{2} \eta^{\mu\rho}(\partial_\nu h_{\rho\sigma} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma}) - \frac{1}{2} h^{\mu\rho}(\partial_\nu h_{\rho\sigma} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma})$$
$$\approx \frac{1}{2} \eta^{\mu\rho}(\partial_\nu h_{\rho\sigma} + \partial_\sigma h_{\nu\rho} - \partial_\rho h_{\nu\sigma}) \text{ to the 1st order}$$

because the $h^{\mu\rho}\,\partial_\alpha h_{\beta\gamma}$ terms are of 2nd order. ✓

A slow-moving particle is one whose velocity components are small compared to $c$:

$$\left| \frac{dx^i}{dt} \right| \ll c, \quad (i = 1, 2, 3). \tag{2.7-5}$$

In flat spacetime, there is no change to the gravitational field. A quasi-static field changes slowly over time. So, we further assume the gravitational field $h_{\mu\nu}$ satisfies the **quasi-static** condition that the rate of change of $h_{\mu\,i}$ with respect to $x^0$ is negligible compared to the rate of change of $h_{\mu\,0}$ with respect to $x^i$. That is, we assume

$$|\,\partial_0 h_{\mu\,i}\,| \ll |\,\partial_i h_{\mu\,0}\,| \ \text{ for all } i \text{ and } \mu. \tag{2.7-6}$$

(The book assumes the quasi-static condition $|\partial_0 h_{\mu\,\nu}| \ll |\partial_i h_{\mu\,\nu}|$ for all $i$, $\mu$, and $\nu$ which is not quite sufficient due to the same index $\nu$ being on both sides of the inequality.) The quasi-static condition is used to obtain equation (2.7-13). The book doesn't provide details, so I provide mine below. Also, see (3.6-4) for additional quasi-static conditions not made clear in the book.

We begin with equation (2.7-1), the geodesic equation of motion for a free particle in spacetime:

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma} \frac{dx^\sigma}{d\tau} \frac{dx^\nu}{d\tau} = 0.$$

We claim that if instead of proper time $\tau$ we use coordinate time $x^0 = ct$, then the geodesic path for a free particle becomes

$$\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma} \frac{dx^\nu}{dt} \frac{dx^\sigma}{dt} = h(t)\, \frac{dx^\mu}{dt} \tag{2.75}$$

where

$$h(t) = -\, \frac{d^2 t}{d\tau^2} \left(\frac{dt}{d\tau}\right)^{-2} = \frac{d^2 \tau}{dt^2} \left(\frac{d\tau}{dt}\right)^{-1} : \tag{2.76}$$

Mimicking Exercise 2.1.1,

$$\frac{d}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt} \qquad \text{(i)}$$

$$\Rightarrow \quad \frac{dx^\mu}{d\tau} = \frac{dt}{d\tau}\frac{dx^\mu}{dt} \qquad \text{(ii)}$$

$$\Rightarrow \quad \frac{d^2 x^\mu}{d\tau^2} = \frac{d}{d\tau}\frac{dx^\mu}{d\tau} \overset{\text{(ii)}}{=} \frac{d}{d\tau}\left(\frac{dt}{d\tau}\frac{dx^\mu}{dt}\right) \overset{\text{(Prdt Rule)}}{=} \frac{dt}{d\tau}\left(\frac{d}{d\tau}\frac{dx^\mu}{dt}\right) + \frac{dx^\mu}{dt}\left(\frac{d}{d\tau}\frac{dt}{d\tau}\right)$$

$$\overset{\text{(i)}}{=} \frac{dt}{d\tau}\frac{dt}{d\tau}\frac{d}{dt}\frac{dx^\mu}{dt} + \frac{dx^\mu}{dt}\frac{d^2 t}{d\tau^2} = \left(\frac{dt}{d\tau}\right)^2 \frac{d^2 x^\mu}{dt^2} + \frac{dx^\mu}{dt}\frac{d^2 t}{d\tau^2} \qquad \text{(iii)}$$

$$\therefore \quad 0 \overset{(2.71)}{=} \frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau}$$

$$\overset{\text{(iii, ii)}}{=} \left(\frac{dt}{d\tau}\right)^2 \frac{d^2 x^\mu}{dt^2} + \frac{dx^\mu}{dt}\frac{d^2 t}{d\tau^2} + \Gamma^\mu_{\nu\sigma}\frac{dt}{d\tau}\frac{dx^\nu}{dt}\frac{dt}{d\tau}\frac{dx^\sigma}{dt}$$

$$= \left(\frac{dt}{d\tau}\right)^2 \left[\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt}\right] + \frac{dx^\mu}{dt}\frac{d^2 t}{d\tau^2}$$

$$= \left(\frac{dt}{d\tau}\right)^2 \left[\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt}\right] + \frac{d^2 t}{d\tau^2}\left[\left(\frac{dt}{d\tau}\right)^{-2}\left(\frac{dt}{d\tau}\right)^2\right]\frac{dx^\mu}{dt}$$

$$= \left(\frac{dt}{d\tau}\right)^2 \left[\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt}\right] - \left[-\frac{d^2 t}{d\tau^2}\left(\frac{dt}{d\tau}\right)^{-2}\right]\left[\left(\frac{dt}{d\tau}\right)^2\frac{dx^\mu}{dt}\right]$$

$$\overset{(2.76)}{=} \left(\frac{dt}{d\tau}\right)^2 \left[\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt} - h(t)\frac{dx^\mu}{dt}\right],$$

or

$$\frac{d^2 x^\mu}{dt^2} + \Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{dt}\frac{dx^\sigma}{dt} = h(t)\frac{dx^\mu}{dt} \quad \text{where} \quad h(t) = -\frac{d^2 t}{d\tau^2}\left(\frac{dt}{d\tau}\right)^{-2} \qquad ✓$$

Next,

$$\frac{d^2 t}{d\tau^2} = \frac{d}{d\tau}\frac{dt}{d\tau} \overset{\text{(i)}}{=} \frac{dt}{d\tau}\frac{d}{dt}\frac{dt}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt}\left(\frac{d\tau}{dt}\right)^{-1} = -\frac{dt}{d\tau}\left(\frac{d\tau}{dt}\right)^{-2}\frac{d^2\tau}{dt^2}$$

$$= -\left(\frac{d\tau}{dt}\right)^{-1}\left(\frac{d\tau}{dt}\right)^{-2}\frac{d^2\tau}{dt^2} = -\left(\frac{d\tau}{dt}\right)^{-3}\frac{d^2\tau}{dt^2} \qquad \text{(iv)}$$

$$\Rightarrow \quad h(t) = -\frac{d^2 t}{d\tau^2}\left(\frac{dt}{d\tau}\right)^{-2} \overset{\text{(iv)}}{=} \left(\frac{d\tau}{dt}\right)^{-3}\frac{d^2\tau}{dt^2}\left(\frac{dt}{d\tau}\right)^{-2} = \left(\frac{d\tau}{dt}\right)^{-3}\frac{d^2\tau}{dt^2}\left(\frac{d\tau}{dt}\right)^2$$

$$= \frac{d^2\tau}{dt^2}\left(\frac{d\tau}{dt}\right)^{-1} \qquad ✓$$

Note. This treatment has a subtle undertone. Until now, when we analyzed the geodesic trajectory, we treated the 4 coordinates as independent, meaning in particular that $\frac{\partial x^i}{\partial t} = 0$ for $i$ = 1, 2, and 3. Here, $\frac{\mathrm{d}x^i}{\mathrm{d}t}$ is not necessarily zero because we are letting the spatial coordinates $x^i$ be functions of $t$. In this development we use (full) derivatives $\frac{\mathrm{d}x^i}{\mathrm{d}t}$, not partial derivatives as we have done heretofore.

The spatial equations of motion (2.75) for a particle along the geodesic path can be written entirely in terms of space coordinates $i$, $j$, and $k$ (i.e., without $\mu$, $\nu$, and $\sigma$) by breaking out the terms with $\nu = 0$ or $\sigma = 0$, yielding:

$$\frac{1}{c^2}\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + \Gamma^i_{00} + 2\,\Gamma^i_{0j}\left(\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) + \Gamma^i_{jk}\left(\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}\right)\left(\frac{1}{c}\frac{\mathrm{d}x^k}{\mathrm{d}t}\right) = \frac{1}{c}h(t)\left(\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}\right): \tag{2.77}$$

First, the three spatial equations (2.75) are written by changing $\mu$ to $i$ :

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + \Gamma^i_{\nu\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}\frac{\mathrm{d}x^\sigma}{\mathrm{d}t} \overset{(2.75)}{=} h(t)\,\frac{\mathrm{d}x^i}{\mathrm{d}t}\ .$$

Observing that $\frac{\mathrm{d}x^0}{\mathrm{d}t} = \frac{\mathrm{d}\,(ct)}{\mathrm{d}t} = c$, equation (2.75) becomes

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + c^2\,\Gamma^i_{00} + 2\,\Gamma^i_{0j}\left(c\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) + \Gamma^i_{jk}\left(\frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\mathrm{d}x^k}{\mathrm{d}t}\right) = h(t)\,\frac{\mathrm{d}x^i}{\mathrm{d}t},$$

or

$$\frac{1}{c^2}\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + \Gamma^i_{00} + 2\,\Gamma^i_{0j}\left(\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) + \Gamma^i_{jk}\left(\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}\right)\left(\frac{1}{c}\frac{\mathrm{d}x^k}{\mathrm{d}t}\right) = \frac{1}{c}\,h(t)\left(\frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t}\right) \quad ✓$$

We next develop Newtonian approximations for the terms in equation (2.77).

The last term on LHS in equation (2.77) is negligible because

$$|\Gamma^i_{jk}| \overset{(2.78,\,\text{above})}{\approx} 0 \text{ to the first order and } \left(\frac{1}{c}\frac{\mathrm{d}x^j}{\mathrm{d}t}\right)\left(\frac{1}{c}\frac{\mathrm{d}x^k}{\mathrm{d}t}\right) \overset{(2.7\text{–}5)}{\ll} 1 \quad . \quad ✓$$

For the 2nd term,

$$\Gamma^i_{00} \overset{(2.78)}{\approx} \frac{1}{2}\,\eta^{i\rho}\,(\partial_0 h_{\rho 0} + \partial_0 h_{0\rho} - \partial_\rho h_{00}) \overset{(2.7\text{–}1)}{=} -\frac{1}{2}\,\delta^\rho_i\,(\partial_0 h_{\rho 0} + \partial_0 h_{0\rho} - \partial_\rho h_{00})$$

$$= -\frac{1}{2}\,(\partial_0 h_{i0} + \partial_0 h_{0i} - \partial_i h_{00}) \overset{(2.7\text{–}6)}{\approx} \frac{1}{2}\,\partial_i h_{00}\ . \tag{2.7-7}$$

Similarly, in the 3rd term,

$$\Gamma^i_{0j} \approx \frac{1}{2}\left(\partial_i h_{0j} - \partial_j h_{0i}\right):$$

$$\Gamma^i_{0j} \overset{(2.78)}{\approx} \frac{1}{2}\eta^{i\rho}\left(\partial_0 h_{j\rho} + \partial_j h_{\rho 0} - \partial_\rho h_{0j}\right) = -\frac{1}{2}\delta^\rho_i\left(\partial_0 h_{j\rho} + \partial_j h_{\rho 0} - \partial_\rho h_{0j}\right)$$

$$= -\frac{1}{2}\left(\partial_0 h_{ji} + \partial_j h_{i0} - \partial_i h_{0j}\right) \overset{(2.7-6)}{\approx} \frac{1}{2}\left(\partial_i h_{0j} - \partial_j h_{0i}\right) \quad ✓$$

As for RHS of equation (2.77), we show that it is negligible:

$$\left(\frac{d\tau}{dt}\right)^2 \overset{(2.69)}{=} \frac{1}{c^2} g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt} = \frac{1}{c^2}\left[ g_{00}\left(\frac{dx^0}{dt}\right)^2 + 2\, g_{0j}\frac{dx^0}{dt}\frac{dx^j}{dt} + g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right]$$

$$= \frac{1}{c^2}\left[ g_{00}\, c^2 + 2\, g_{0j}\, c\frac{dx^j}{dt} + g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right].$$

Since $2\,|\, g_{0j}\, c\,\frac{dx^j}{dt}| \overset{(2.7-5)}{\ll} c^2 \approx g_{00}c^2$, and $|\, g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\,| \overset{(2.7-5)}{\ll} c^2 \approx g_{00}c^2$, then

$$\left(\frac{d\tau}{dt}\right)^2 \approx \frac{1}{c^2}\left[ g_{00}\, c^2\right] = g_{00} = \eta_{00} + h_{00} = 1 + h_{00}$$

$$\frac{d\tau}{dt} \approx \sqrt{1+h_{00}} \approx 1 + \frac{1}{2}h_{00} \tag{2.79}$$

$$\left(\frac{d\tau}{dt}\right)^{-1} \overset{(2.79)}{\approx} \frac{1}{1+\frac{1}{2}h_{00}}\;\frac{1-\frac{1}{2}h_{00}}{1-\frac{1}{2}h_{00}} = \frac{1-\frac{1}{2}h_{00}}{1-\frac{1}{4}{h_{00}}^2} \approx 1 - \frac{1}{2}h_{00} \tag{2.7-8}$$

$$\frac{d^2\tau}{dt^2} \overset{(2.79)}{\approx} \frac{1}{2}\frac{dh_{00}}{dt} = \frac{c}{2}\frac{dh_{00}}{d(c\,t)} = \frac{c}{2}\frac{dh_{00}}{dx^0} = \frac{c}{2}\, h_{00,0} \tag{2.7-9}$$

$$\frac{1}{c}\, h(t) \overset{(2.76)}{=} \frac{1}{c}\frac{d^2\tau}{dt^2}\left(\frac{d\tau}{dt}\right)^{-1} \overset{(2.7-9)}{=} \frac{1}{2}h_{00,0}\left(\frac{d\tau}{dt}\right)^{-1} \overset{(2.7-8)}{\approx} \frac{1}{2}h_{00,0}\left(1 - \frac{1}{2}h_{00}\right) \tag{2.7-10}$$

Because $h_{00} \overset{(2.7-2)}{\approx} 0$,

$$1 - \frac{1}{2}\, h_{00} \approx 1\,. \tag{2.7-11}$$

Also, $\frac{1}{c}\left|\frac{dx^i}{dt}\right| \overset{(2.7-5)}{\ll} 1.$ (2.7-12)

Although $h_{00}$ is very small, $|\, h_{00,0} \,|$ can in general be large. So, we impose an additional quasi-static condition that $|\, h_{00,0} \,| < 1$. Then,

RHS of equation (2.77) ≈ 0:

$$|\mathrm{RHS}| \overset{(2.77)}{=} \frac{1}{c}\,|h(t)|\,\frac{1}{c}\,\left|\frac{\mathrm{d}x^i}{\mathrm{d}t}\right| \overset{(2.7-10)}{\approx} \frac{1}{2}\,|h_{00,0}|\,\left|1-\frac{1}{2}\,h_{00}\right|\,\frac{1}{c}\,\left|\frac{\mathrm{d}x^i}{\mathrm{d}t}\right|$$
$$\overset{(2.7-11,\,2.7-12)}{\ll} |\, h_{00,0} \,| \approx 0 \qquad ✓$$

Thus, equation (2.77) for the geodesic path of a slow-moving free particle simplifies to

$$\frac{1}{c^2}\,\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} + \tfrac{1}{2}\,\partial_i\, h_{00} + (\partial_i h_{0j} - \partial_j h_{0i})\left(\frac{1}{c}\,\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) = 0. \qquad (2.7\text{-}13)$$

Introducing the mass $m$ of the particle and rearranging yields

$$m\,\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} \overset{(2.7-13)}{=} -m\,\partial_i\left(\frac{c^2}{2}\,h_{00}\right) + mc\,(\partial_i h_{0j} - \partial_j h_{0i})\left(\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) \qquad (2.80)$$

In Newtonian terms we can interpret LHS as mass × acceleration, $m\,\frac{\mathrm{d}^2 \boldsymbol{x}}{\mathrm{d}t^2}$; i.e., the "gravitational force" on the particle. The first term on RHS is the force $-m\,\nabla V$ arising from a potential $V = \frac{c^2}{2}\,h_{00}$. The second term on RHS is velocity-dependent and has the feel of a rotation:

Recall the rotation matrix of Appendix A.1. In the rotation matrix,
$M_{ij} = -\,M_{ji}$ (= $sin\theta$), similar to $M_{ij} = \partial_i h_{0j} - \partial_j h_{0i} = -\,(\partial_j h_{0i} - \partial_i h_{0j}) = -\,M_{ji}$.)

This is consistent with the principle of equivalence that asserts that forces of acceleration, such as the velocity-dependent Coriolis force which would arise from using a rotating reference system, are on the same footing as the gravitational forces. If we agree to call nearly inertial coordinate system in which $\partial_i h_{0j} - \partial_j h_{0i} = 0$ **non-rotating**, then equation (2.80) reduces to

$$\frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = -\,\partial_i V \qquad (2.81)$$

for slowly moving particles (see equation (2.75), where

$$V = \frac{c^2}{2}\,h_{00}\,. \qquad (2.82)$$

Equation (2.81) is the Newtonian equation of motion, equation (2.6-2), for a particle moving in a gravitational field of potential $V$. From equation (2.82) we get

$$g_{00} = 1 + \frac{2\,V}{c^2}: \tag{2.83}$$

$$g_{00} \overset{(2.74)}{=} \eta_{00} + h_{00} \overset{(2.75)}{=} 1 + h_{00} \overset{(2.82)}{=} 1 + \frac{2\,V}{c^2} \qquad \checkmark$$

Equation (2.83) is the relation we have been seeking between $g_{00}$ and the acceleration potential $V$ in Newton’s 2nd Law, equation (2.6-2), the Newtonian approximation to the geodesic equation.

## 2.8 Newton's law of universal gravitation

In Appendix A, Section 6, Newton's second law in Special Relativity was given in equation (A.29) as

$$f^\mu = \frac{dp^\mu}{d\tau}.$$

Using the rules for changing Special Relativity tensors into General Relativity tensors, the second law becomes

$$f^\mu = \frac{Dp^\mu}{d\tau}. \qquad (2.8\text{-}1)$$

As a test of the development of Einstein's field equations, we wish to show that for flat Euclidean 3-space, the field equations generate ***F*** = *m**a,*** where *F* satisfies Newton's equation of universal gravitation (2.6-5), $F = \frac{G\,M\,m}{r^2}$.

The **Schwarzschild solution**, which will be derived in Section 3.7, is an exact solution of the field equations of general relativity. Its line element is

$$c^2\,d\tau^2 = (1 - \frac{2\,GM}{r\,c^2})\,c^2 dt^2 - \left(1 - \frac{2\,GM}{r\,c^2}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 sin^2\theta\,d\phi^2 \qquad (2.8\text{-}2)$$

where $M$ is the mass of the body and $G$ is the gravitational constant. Newton's law of gravitation for Euclidean 3-space can be shown to generated from this as follows.

For small values of $\frac{GM}{r\,c^2}$, line element (2.8-2) is approximates the Euclidean flat spacetime line element in spherical coordinates, equation (1.6-9):

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 sin^2\theta\,d\phi^2.$$

Also, for small values of $\frac{GM}{r\,c^2}$, we see that $g_{00} \approx 1$. When we write the metric tensor in the form $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, then $g_{00} = 1 + h_{00}$, so $h_{00} \overset{(2.8-2)}{=} -\frac{2\,GM}{r\,c^2}$ is small.

For slowly moving particles, a gravitational potential, $V$, can be defined:

$$V \overset{(2.82)}{=} \frac{c^2}{2}\,h_{00} = -\frac{GM}{r}. \qquad (2.8\text{-}3)$$

Then,

$$\frac{d^2 x^i}{dt^2} \overset{(2.81)}{=} -\partial_i V \overset{(2.8-3)}{=} GM \frac{\partial}{\partial x^i}(r^{-1}) = -\frac{GM}{r^2}\frac{dr}{dx^i}\ . \tag{2.8-4}$$

Letting $\boldsymbol{r} = (x^1, x^2, x^3)$, and letting $\hat{\boldsymbol{r}} = (\frac{dr}{dx^1}, \frac{dr}{dx^2}, \frac{dr}{dx^3})$, the unit vector in the direction of $\boldsymbol{r}$, we get

$$\boldsymbol{F} = m\boldsymbol{a} = m\frac{d^2 \boldsymbol{r}}{dt^2} = m\left(\frac{d^2 x^1}{dt^2}, \frac{d^2 x^2}{dt^2}, \frac{d^2 x^3}{dt^2}\right) \overset{(2.8-4)}{=} -\frac{GMm}{r^2}\left(\frac{dr}{dx^1}, \frac{dr}{dx^2}, \frac{dr}{dx^3}\right)$$
$$= -\frac{GMm}{r^2}\hat{\boldsymbol{r}}.$$

The magnitude of the force is $F = \frac{G M m}{r^2}$, which is Newton's equation for universal gravitation.

We have shown, using the Schwarzschild exact solution of Einstein's field equations, that Newton's second law holds for small values of $\frac{GM}{r c^2}$ and for particles traveling at non-relativistic speeds.

## 2.9 A rotating reference system

In a reference system that is rotating with a non-constant rotational velocity, Newton's equations of motion fail to match observation until 3 fictional forces are included.

| | |
|---|---|
| **Euler Force** | $-m\frac{d\boldsymbol{\omega}}{\mathrm{d}t}\times\boldsymbol{r}'$ |
| **Coriolis Force** | $-2m\,(\boldsymbol{\omega}\times\boldsymbol{v}')$ |
| **Centrifugal Force** | $-m\,\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r}')$ |

where the primed reference system is a rotating coordinate system, and
$\boldsymbol{\omega}$ is the angular velocity of the rotating frame relative to the inertial system,
$\boldsymbol{v}'$ is the velocity of an object in the primed system,
$\boldsymbol{r}'$ is the position vector of an object in the primed system,
$m$ is the mass of the object.

On a merry-go-round, the Euler force is felt as a force that pushes you backward as the merry-go-round starts up, and it pulls you forward as it decelerates.

Consider a rotating Earth with the origin of both coordinate systems on the equator. Suppose a car travels directly to the North Pole at constant speed. In the inertial system, there are no forces acting on the car, but from the point of view of an observer on the rotating equator, the car appears subject to the Coriolis force that curves the path of the car to the left. In general relativity this force is accounted for by treating the path as a geodesic. It is on the same footing as the fictional gravitational force.

The centrifugal force is a radially outward force. If we imagine a string attached to a swinging ball, the ball exerts a centrifugal force on the string, and the string exerts and equal but opposite centripetal force on the ball. Again, general relativity doesn't recognize this as a force but rather handles it as a geodesic path.

If the frame rotates at a constant angular velocity, the Euler force disappears. If the particle is located at the origin of the primed system, the centrifugal force disappears. If the angular velocity is zero, all three fictional forces disappear.

All the imaginary forces, including gravity, are interwoven into the connection coefficients. Given connection coefficients, it can be from difficult to impossible to break out the contributing imaginary forces. However, we can identify them in the case of a simple rotating Cartesian system rotating at a constant angular velocity, as is now demonstrated.

Let K be an inertial (non-rotating) system with coordinates $(T,X,Y,Z)$ and line element

$$c^2\,d\tau^2 = c^2\,dT^2 - dX^2 - dY^2 - dZ^2\,. \tag{2.84}$$

Denote $X^0 = cT$, $X^1 = X$, $X^2 = Y$, $X^3 = Z$.

Define a rotating coordinate system K′ having coordinates $(t,x,y,z)$ implicitly defined by

$$\begin{aligned} T &= t \\ X &= x\cos\omega t - y\sin\omega t \\ Y &= x\sin\omega t + y\cos\omega t \\ Z &= z\,. \end{aligned} \tag{2.85}$$

Denote $x^0 = ct$, $x^1 = x$, $x^2 = y$, $x^3 = z$.

Points $(x,y,z)$ in K′ rotate counter-clockwise about the $Z$-axis of K with constant angular velocity $\omega$.

Using "dot" to denote the derivative with respect to $\tau$, we show in Exercise 2.9.1 that the line element in K′ is

$$c^2\,d\tau^2 = [c^2 - \omega^2(x^2 + y^2)]\,dt^2 + 2\omega y\,dx\,dt - 2\omega x\,dy\,dt - dx^2 - dy^2 - dz^2 \tag{2.86}$$

We then show in Exercise 2.9.2 that the spacetime geodesic equations (2.71) for a free particle with mass, in the rotating frame K′, are

$$\begin{aligned} \ddot{t} &= 0 \\ \ddot{x} - \omega^2 x\,\dot{t}^2 - 2\omega\,\dot{y}\,\dot{t} &= 0 \\ \ddot{y} - \omega^2 y\,\dot{t}^2 + 2\omega\,\dot{x}\,\dot{t} &= 0 \\ \ddot{z} &= 0\,. \end{aligned} \tag{2.87}$$

These constitute the equations of motion of a free particle (with mass).

In Exercise 2.9.2, we are actually asked to do more, to derive equation (2.87) using three different methods:

(a) Use the Euler-Lagrange equations (2.17): $\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^\sigma}\right) - \frac{\partial L}{\partial x^\sigma} = 0$, where the Lagrangian is $L \overset{(2.1\text{–}5)}{=} \frac{1}{2} g_{\mu\nu}\, \dot{x}^\mu\, \dot{x}^\gamma$. (The Euler-Lagrange equations are equivalent to the geodesic equations, so the approach is to extract $g_{\mu\nu}$ from the line element (2.86), then compute $\frac{\partial L}{\partial x^\sigma}$, $\frac{\partial L}{\partial \dot{x}^\sigma}$, and $\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^\sigma}\right)$ for $\sigma = 0 - 3$.)

(b) Calculate $g^{\mu\nu}$, the inverse of $g_{\mu\nu}$ obtained (a), then calculate $\Gamma^\mu_{\nu\sigma} = g^{\mu\delta}\Gamma_{\delta\nu\gamma}$, and, lastly, write out the four geodesic equations (2.87): $\ddot{x} + \Gamma^\mu_{\nu\gamma}\dot{x}^\nu\dot{x}^\gamma = 0$.

(c) Show that $\ddot{T} = \ddot{X} = \ddot{Y} = \ddot{Z} = 0$, and then compute these quantities using equations (2.85).

We are now prepared to unravel the forces contained in the connection coefficients. Instead of using the usual primed coordinates we stick with the simpler lower-case notation of equation (2.87) for the primed system. The first of equations (2.87) implies $\dot{t} = \frac{dt}{d\tau} = k$, a constant. So

$$\dot{x}^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau} = k\,\frac{dx^\mu}{dt} \;\Rightarrow\; \ddot{x}^\mu = \frac{d^2 x^\mu}{d\tau^2} = \frac{d^2 x^\mu}{dt^2}\left(\frac{dt}{d\tau}\right)^2 = k^2\,\frac{d^2 x^\mu}{dt^2}.$$

Plugging this into the three space equations of (2.87), and factoring out $k^2$, yields

$$\frac{d^2 x}{dt^2} - \omega^2 x - 2\,\omega\,\frac{dy}{dt} = 0$$
$$\frac{d^2 y}{dt^2} - \omega^2 y + 2\,\omega\,\frac{dx}{dt} = 0$$
$$\frac{d^2 z}{dt^2} = 0\,.$$

Introducing mass and rearranging yields

$$\begin{aligned} m\frac{d^2 x}{dt^2} &= m\omega^2 x + 2m\omega\,\frac{dy}{dt} \\ m\frac{d^2 y}{dt^2} &= m\omega^2 y - 2m\omega\,\frac{dx}{dt} \\ m\frac{d^2 z}{dt^2} &= 0\,. \end{aligned} \tag{2.88}$$

With $\boldsymbol{r}$ = x $\boldsymbol{i}$ + y $\boldsymbol{j}$ + z $\boldsymbol{k}$, and setting $\boldsymbol{\omega} = \omega\,\boldsymbol{k}$, this can be written as a vector equation:

$$m\frac{d^2\boldsymbol{r}}{dt^2} = -\,m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r}) - 2m\boldsymbol{\omega}\times\frac{d\boldsymbol{r}}{dt} : \tag{2.89}$$

$$\frac{d^2\boldsymbol{r}}{dt^2} = \frac{d^2x}{dt^2}\,\boldsymbol{i} + \frac{d^2y}{dt^2}\,\boldsymbol{j} + \frac{d^2z}{dt^2}\,\boldsymbol{k} = (\omega^2x + 2\omega\frac{dy}{dt})\,\boldsymbol{i} + (\omega^2y - 2\omega\frac{dx}{dt})\,\boldsymbol{j} + 0\,\boldsymbol{k}$$

$$\boldsymbol{\omega}\times\boldsymbol{r} = (\omega_y z - \omega_z y)\,\boldsymbol{i} + (\omega_z x - \omega_x z)\,\boldsymbol{j} + (\omega_x y - \omega_y x)\,\boldsymbol{k} = -\,\omega y\,\boldsymbol{i} + \omega x\,\boldsymbol{j}$$

$$\boldsymbol{\omega}\times\frac{d\boldsymbol{r}}{dt} = -\,\omega\frac{dy}{dt}\,\boldsymbol{i} + \omega\frac{dx}{dt}\,\boldsymbol{j}$$

$$\begin{aligned}\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r}) &= [\omega_y(0) - \omega_z(\omega x)]\,\boldsymbol{i} + [\omega_z(-\omega y) - \omega_x(0)]\,\boldsymbol{j} + [\omega_x(\omega x) - \omega_y(-\omega y)]\,\boldsymbol{k}\\ &= -\,\omega^2 x\,\boldsymbol{i} - \omega^2 y\,\boldsymbol{j}\end{aligned}$$

Thus,

$$-\,\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r}) - 2\boldsymbol{\omega}\times\frac{d\boldsymbol{r}}{dt} = (\omega^2x + 2\omega\frac{dy}{dt})\,\boldsymbol{i} + (\omega^2y - 2\omega\frac{dx}{dt})\,\boldsymbol{j} = \frac{d^2\boldsymbol{r}}{dt^2} \quad ✓$$

Equation (2.89) unravels the forces that are intertwined in the connection coefficients. We recognize the centrifugal force, $-\,m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r})$, and the Coriolis force, $-\,2m\boldsymbol{\omega}\times\frac{d\boldsymbol{r}}{dt}$ on the RHS as the two pieces of the total force, $m\boldsymbol{a}$, on the LHS.

An alternate way to derive equation (2.89) is to use the approximation methods of Section 2.7, by noting that

$$(g_{\mu\nu}) = \begin{pmatrix} 1-\frac{(x^2+y^2)\,\omega^2}{c^2} & \frac{y\,\omega}{c} & -\frac{x\,\omega}{c} & 0\\ \frac{y\,\omega}{c} & -1 & 0 & 0\\ -\frac{x\,\omega}{c} & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{pmatrix}$$

$$= \begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} + \begin{pmatrix} -\frac{(x^2+y^2)\,\omega^2}{c^2} & \frac{y\,\omega}{c} & -\frac{x\,\omega}{c} & 0\\ \frac{y\,\omega}{c} & 0 & 0 & 0\\ -\frac{x\,\omega}{c} & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

$$\equiv (\eta_{\mu\nu}) + (h_{\mu\nu}),$$

and that the $h_{\mu\nu}$ are small near the $z$-axis because $x^2 + y^2$ are small. From the matrix above, the quasi-static condition $\partial_0 h_{\mu\nu} = 0$ holds and, so, equation (2.80) applies, from which the equations (2.89) can be (eventually) obtained.

# Chapter 3 Field Equations and Curvature

## 3.0 Introduction

The General Relativity (GR) field equations quantify the relationship between energy and space curvature. In flat Euclidean space, there is a single field equation, Poisson's equation (2.6-11):

$$\nabla^2 V = 4\pi G\rho.$$

This is a field equation because LHS contains the energy potential $V$, developed in equation (2.6-2), which represents all possible sources of energy, and RHS contains $G$, which represents gravity, the source of curvature in general relativity.

In Section 3.1 we introduce the  tensor, $T^{\mu\nu}$, which represents all sources of energy. We show the relationship between $T^{\mu\nu}$ and the relativistic continuity equation and equation of motion, analogous to Section 2.6 where we developed Poisson's equation and the fluid continuity equation from Newton's equation of motion.

In Sections 3.2 – 3.4, we develop curvature, including the various curvature tensors, parallel transport, and geodesic deviation.

In Section 3.5 we present the GR field equations, and in Section 3.6 we show that they meet the constraint that Poisson's equation be satisfied in flat Euclidean space at non-relativistic speeds.

The Schwarzschild exact solution is derived in Section 3.7.

## 3.1 The stress tensor and fluid motion

We begin in *flat spacetime* using the ($ct$,$x$,$y$,$z$) inertial coordinate system, then extend the equations to curved spacetime, and then to non-inertial curved spacetime.

Convention  Boldface is used for 3-vectors, and non boldface for 4-vectors:

$$\lambda^\mu \equiv (\lambda^0, \lambda^1, \lambda^2, \lambda^3) \equiv (\lambda^0, \boldsymbol{\lambda}) .$$

We start with a particle and use the following notation:

$m \equiv$ **rest or proper mass of particle**

$t \equiv$ **coordinate time**

$\tau \equiv$ **proper time**

$v \equiv$ **speed of particle**

$$\gamma \equiv \frac{\mathrm{d}t}{\mathrm{d}\tau} \overset{(A.42)}{=} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

$E \equiv \gamma mc^2 \equiv$ **energy of particle** (3.1-1)

$v^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}t} \equiv$ **coordinate velocity of particle**

$u^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \gamma v^\mu \equiv$ **world velocity of particle** (3.1-2)

$p^\mu \equiv mu^\mu \equiv$ **4-momentum of particle** (3.1-3)

Since $\dfrac{\mathrm{d}x^0}{\mathrm{d}t} = \dfrac{\mathrm{d}(ct)}{\mathrm{d}t} = c$, by the convention above we get that

$$v^\mu = (c,\mathbf{v}) \tag{3.1-4}$$

where $\mathbf{v} = (v^1, v^2, v^3)$ is the velocity 3-vector. Hence, the particle's speed is

$$v = |\mathbf{v}|, \text{ and } v^2 = (v^1)^2 + (v^2)^2 + (v^3)^2 \tag{3.1-5}$$

Of the quantities listed above, only $m = m(x^\mu)$ and $\tau = \tau(x^\mu)$ are scalars and only $u^\mu$ and $p^\mu$ are vectors:

We showed in Section A.6 that,$v^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}t}$ is not a vector because $\dfrac{\mathrm{d}\,x^\mu}{\mathrm{d}t'} \neq \dfrac{\mathrm{d}x^\mu}{\mathrm{d}t}$,

whereas $u^\mu \equiv \dfrac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ *is* as vector because $\tau' = \tau$.

$E$, $v$, and $t$ are not functions of ($x^\mu$) so, by definition (1.7-4), they are not scalars.

For a moving particle,

$$u^\mu \overset{(3.1-2)}{=} \gamma v^\mu \overset{(3.1-4)}{=} \gamma(c,\boldsymbol{v}) \qquad (3.1\text{-}6)$$

$$p^\mu \overset{(3.1-3)}{=} m u^\mu \overset{(3.1-6)}{=} m\gamma(c,\boldsymbol{v}) = (m\gamma c, m\gamma \boldsymbol{v}) \overset{(3.1-1)}{=} \left(\frac{E}{c}, \boldsymbol{p}\right) \qquad (3.1)$$

where the 3-momentum $\boldsymbol{p} \equiv m\gamma\boldsymbol{v}$. (3.1-7)

For a stationary particle, $\gamma = 1$, $\boldsymbol{v} = 0$, and so

$$u^\mu \overset{(3.1-6)}{=} \gamma(c,\boldsymbol{v}) = (c,\boldsymbol{0}) \qquad (3.1\text{-}8)$$

$$p^\mu \overset{(3.1-3)}{=} m u^\mu = m(c,\boldsymbol{0})$$

$p^0 = mc = \dfrac{mc^2}{c} \overset{(3.1-1)}{=} \dfrac{E}{c}$ is proportional to the **rest energy**, $E \equiv mc^2$

Equation (3.1) reminds us that energy and 3-momentum are components of a single momentum 4-vector, just as time and position are components of the position 4-vector.

Another useful formula is

$$u^\mu u_\mu = c^2 : \qquad (3.1\text{-}9)$$

$$u^\mu u_\mu \overset{(3.1-6)}{=} \gamma^2\, v^\mu v_\mu \overset{(\text{Example } A.0.1\,(\text{b}))}{=} \gamma^2\, [(v^0)^2 - (v^1)^2 - (v^2)^2 - (v^3)^2]$$

$$\overset{(3.1-4,\, 3.1-5)}{=} \gamma^2\,(c^2 - v^2) \overset{(A.42)}{=} \frac{c^2 - v^2}{1 - \frac{v^2}{c^2}} = c^2\ ✓$$

We next extend these concepts to the continuous case. The simplest such environment is a perfect fluid characterized by two scalar fields, density $\rho$ and pressure $P$, and a vector field, world velocity $u^\mu$. In order for $\rho$ to be a scalar field, we must define it to be **proper density**, the rest mass per unit rest volume. In calculations we now use the **4-momentum density** $\rho u^\mu$ in place of the particle 4-momentum $p^\mu = mu^\mu$. We italicize pressure $P$ in order to distinguish it from a point P, although the context should always make the distinction clear. (The book uses *p,* which looks a lot like $\rho$, and is also the symbol that should represent the magnitude of the 3-momentum, $\boldsymbol{p}$.)

For a particle, $p^\mu$ is a tensor that captures both energy and momentum. For a fluid, we seek a tensor that captures the energy and 4-momentum density of the fluid. The derivation is beyond the scope of this book so we just give the tensor and motivate it by describing its properties.

Definition The **energy-momentum-stress tensor** (or, **stress tensor** for short) for a perfect fluid is defined as

$$\boxed{T^{\mu\nu} \equiv \left(\rho + \frac{P}{c^2}\right) u^\mu u^\nu - P\eta^{\mu\nu}} \quad , \tag{3.2}$$

where $\eta$ is the metric tensor for special relativity defined in equation (A.4).

Observe that $T^{\mu\nu}$ is symmetric and is composed of $\rho$, $P$, and $u^\mu$, the scalar and vector fields that characterize the fluid. Next, observe that

$$T^{\mu\nu} u_\nu = c^2 \rho u^\mu : \tag{3.1-10}$$

$$T^{\mu\nu} u_\nu \overset{(3.2)}{=} (\rho + \frac{P}{c^2}) u^\mu u^\nu u_\nu - P\, \eta^{\mu\nu} u_\nu \overset{(3.1-9)}{=} (\rho + \frac{P}{c^2}) u^\mu c^2 - P u^\mu = c^2 \rho u^\mu \quad ✓$$

That is, $T^{\mu\nu} u_\nu$ equals the 4-momentum density $\rho u^\mu$ up to a factor of $c^2$.

Recalling that $\lambda^\nu{}_{,\mu} = \dfrac{\partial \lambda^\nu}{\partial x^\mu}$ so that, for example, $u^\mu{}_{,\mu} = \dfrac{\partial u^\mu}{\partial x^\mu} = \dfrac{\partial u^0}{\partial x^0} + \dfrac{\partial u^1}{\partial x^1} + \dfrac{\partial u^2}{\partial x^2} + \dfrac{\partial u^3}{\partial x^3}$, we have the following definitions.

Definition The **relativistic continuity equation** is

$$\boxed{(\rho u^\mu)_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu} = 0} \tag{3.5}$$

and the **equation of motion of a perfect fluid** is

$$\boxed{\left(\rho + \frac{P}{c^2}\right) u^\nu{}_{,\mu} u^\mu = \left(\eta^{\mu\nu} - \frac{1}{c^2} u^\mu u^\nu\right) P_{,\mu}} \, . \tag{3.6}$$

Recall that for 3-vectors, **divergence** was defined as

$$\textbf{div}\ \boldsymbol{v} \overset{(2.61\,A)}{\equiv} \nabla \bullet \boldsymbol{v} = \frac{\partial v^i}{\partial x^i} = v^i{}_{,i} \, .$$

Since the stress tensor is symmetric, it also only has one **divergence:**

$$\textbf{div}\ \boldsymbol{T}^{\mu\nu} \equiv \nabla \bullet T^{\mu\nu} \overset{(2.61\,D)}{\equiv} T^{\mu\nu}{}_{,\mu} \, . \tag{3.1-11}$$

Theorem 3.1.1 The divergence $T^{\mu\nu}{}_{,\mu} = 0$ iff equations (3.5) and (3.6) hold.

Proof. First, suppose $T^{\mu\nu}{}_{,\mu} = 0$. Since $u^\nu u_\nu \overset{(3.1-9)}{=} c^2$ we get

$$u^\nu{}_{,\mu}\, u_\nu + u^\nu\, u_{\nu,\mu} = 0\,. \qquad (3.4)$$

Therefore (Exercise 3.1.4) $u^\nu{}_{,\mu}\, u_\nu = 0$ : (3.1-12)

$$\begin{aligned} 0 &\overset{(3.4)}{=} u^\nu{}_{,\mu}\, u_\nu + u^\nu\, u_{\nu,\mu} \\ &\overset{(\text{Ex A0.1 a})}{=} (u^0 u^0{}_{,\mu} - u^1 u^1{}_{,\mu} - u^2 u^2{}_{,\mu} - u^3 u^3{}_{,\mu}) + (u^0 u^0{}_{,\mu} - u^1 u^1{}_{,\mu} - u^2 u^2{}_{,\mu} - u^3 u^3{}_{,\mu}) \\ &= 2\, u^\nu{}_{,\mu}\, u_\mu \quad ✓ \end{aligned}$$

Differentiating equation (3.2) with respect to $x^\mu$ yields

$$0 = T^{\mu\nu}{}_{,\mu} = \rho u^\mu{}_{,\mu} u^\nu + \rho u^\mu u^\nu{}_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu} u^\nu + \frac{P}{c^2} u^\mu u^\nu{}_{,\mu} + \frac{1}{c^2} P_{,\mu} u^\mu u^\nu - P_{,\mu} \eta^{\mu\nu} \qquad (3.3)$$

Contracting equation (3.3) with $u_\nu$ and applying (3.1-9) yields equation (3.5) :

$$\begin{aligned} 0 &= \rho\, u^\mu{}_{,\mu}\, c^2 + \rho\, u^\nu{}_{,\mu}\, u^\mu\, u_\nu + \frac{P}{c^2} u^\mu{}_{,\mu}\, c^2 + \frac{P}{c^2} u^\mu\, u_\nu\, u^\nu{}_{,\mu} + \frac{1}{c^2} P_{,\mu}\, u^\mu\, c^2 - P_{,\mu}\, u_\nu\, \eta^{\mu\nu} \\ &= c^2\, (\rho\, u^\mu{}_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu}) + u^\mu\, u_\nu\, u^\nu{}_{,\mu}\, (\rho + \frac{P}{c^2}) + P_{,\mu}\, (u^\mu - u^\mu) \\ &\overset{(3.4)}{=} c^2\, (\rho\, u^\mu{}_{,\mu} + \frac{P}{c^2} u^\mu{}_{,\mu}) \qquad ✓ \end{aligned}$$

Applying equation (3.5) to equation (3.3) yields equation (3.6). ✓
Conversely, if equations (3.5) and (3.6) hold, plugging them into equation (3.2) yields $T^{\mu\nu}{}_{,\mu} = 0$. ■

In an instantaneous rest system the coordinate divergence is zero, so the equation $T^{\mu\nu}{}_{,\mu} = 0$ is equivalent to setting its zeroth component to zero: $T^{\mu 0}{}_{,\mu} = 0$ :

For a particle at rest,

$$u^\mu \overset{(3.1-8)}{=} (c,\mathbf{0}) \;\Rightarrow\; u_\nu \overset{(\text{Ex } A.0\times.1)}{=} u_0 = c \text{ and } u_i = 0 \text{ for } i = 1,2,3.$$

So,

$$0 = T^{\mu\nu}{}_{,\mu}\, u_\nu = T^{\mu 0}{}_{,\mu}\, u_o + T^{\mu i}{}_{,\mu}\, u_i = c\, T^{\mu 0}{}_{,\mu} \;\Leftrightarrow\; T^{\mu 0}{}_{,\mu} = 0 \quad ✓$$

Equations (3.7) and (3.9), below, were derived in Section 2.6 as equation (2.6-12) and (2.6-15), respectively. The first represents conservation of mass in fluid flow; the second is the fluid equation of motion, $\boldsymbol{F} = m\boldsymbol{a}$, where force on a fluid is represented by pressure and mass is represented by density, and it and includes both internal and external forces.

Definition The **classical continuity equation for a perfect fluid** is

$$\boxed{\partial\rho/\partial t + \nabla\cdot(\rho\mathbf{v}) = 0} \qquad (3.7)$$

where (in the classical case) $\rho$ is both the proper density (i.e., global flat spacetime coordinates) and the coordinate density (i.e., local coordinates), and

$$\nabla\cdot(\rho\mathbf{v}) = \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right)\cdot(\rho v^1\mathbf{i} + \rho v^2\mathbf{j} + \rho v^3\mathbf{k}) = \frac{\partial}{\partial x^i}(\rho\, v^i) \qquad (3.1\text{-}13)$$

Euler's **classical equation of motion for a perfect fluid** is

$$\boxed{\rho\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right)\mathbf{v} = -\nabla P} \qquad (3.9)$$

where

$$\mathbf{v}\cdot\nabla = v^i\frac{\partial}{\partial x^i}. \qquad (3.1\text{-}14)$$

We now show that the relativistic continuity and motion equations (3.5) and (3.6) reduce, respectively, to equations (3.7) and (3.9) in the classical limit of a slowly moving fluid ($v \ll c$) and small pressure ($P \ll c^2$). In this case,

$$\gamma = 1,\ \ u^\mu = v^\mu,\ \ \frac{v^0}{c} = 1,\ \ \frac{v^i}{c} = 0,\ \text{ and }\ \frac{P}{c^2} = 0. \qquad (3.1\text{-}15)$$

Equation (3.5), $(\rho\, u^\mu)_{,\mu} + \frac{P}{c^2}u^\mu{}_{,\mu} = 0$, reduces to

$$(\rho\, c)_{,0} + \left(\rho\, v^i\right)_{,i} = 0: \qquad (3.1\text{-}16)$$

$$(\rho\, u^\mu)_{,\mu} + \frac{P}{c^2}u^\mu{}_{,\mu} \overset{(3.1-15)}{=} (\rho\, u^\mu)_{,\mu} \overset{(3.1-6)}{=} (\rho\,\gamma\, c)_{,0} + \left(\rho\ \gamma\, v^i\right)_{,i}$$

$$\overset{(3.1-15)}{=} (\rho\, c)_{,0} + \left(\rho\ v^i\right)_{,i}\ \ ✓$$

which in 3-vector notation is equation (3.7):

$$0 \overset{(3.1-16)}{=} \frac{\partial(\rho\, c)}{\partial(c\, t)} + \frac{\partial}{\partial x^i}(\rho\, v^i) \overset{(3.1-13)}{=} \frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) \qquad ✓ \qquad ■$$

Equation (3.6), $(\rho + \frac{P}{c^2})\, u^\nu{}_{,\mu}\, u^\mu = (\eta^{\mu\nu} - \frac{1}{c^2} u^\mu u^\nu)\, P_{,\mu}$, reduces to

$$\rho\, v^\nu{}_{,\mu} v^\mu = (\eta^{\mu\nu} - \frac{1}{c^2} v^\mu v^\nu)\, P_{,\mu} \tag{3.8}$$

because $\frac{P}{c^2} \overset{(3.1\text{–}15)}{=} 0$ and $u^\mu \overset{(3.1\text{–}15)}{=} v^\mu$.

The zeroth component of equation (3.8) is 0 = 0:

$$\text{LHS} = \rho\, v^0{}_{,\mu} v^\mu = \rho \frac{\partial v^0}{\partial x^\mu} v^\mu \overset{(3.1\text{–}4)}{=} \rho \frac{\partial c}{\partial x^\mu} v^\mu = 0 \quad ✓$$

$$\text{RHS} = (\eta^{\mu 0} - \frac{1}{c^2} v^\mu v^0)\, P_{,\mu} = [\eta^{00} - \frac{1}{c^2}(v^0)^2\,]\, P_{,0} + [\eta^{i\,0} - \frac{1}{c^2} v^i v^0]\;\; P_{,i}$$

$$\overset{(3.1\text{–}4)}{=} [1 - \frac{c^2}{c^2}]\, P_{,0} + [0 - \frac{1}{c^2} v^i v^0]\;\; P_{,i} \overset{(3.1\text{–}15)}{=} 0 + [0 - 0] = 0 \quad ✓$$

The non-zero components are

$$\rho\, v^i{}_{,\mu} v^\mu \overset{(3.8)}{=} (\eta^{\mu\, i} - \frac{1}{c^2} v^\mu v^i)\, P_{,\mu} = (\eta^{0\, i} - \frac{1}{c^2} v^0 v^i)\, P_{,0} + (\eta^{j\, i} - \frac{1}{c^2} v^j v^i)\, P_{,j}$$

$$\overset{(3.1\text{–}4)}{=} (0 - \frac{v^i}{c})\, P_{,0} + (-\,\delta^{j\,i} - \frac{1}{c^2} v^j v^i)\, P_{,j} \overset{(3.1\text{–}15)}{=} (0\text{–}\, 0)\, P_{,0} - \delta^{j\,i}\, P_{,j}$$

$$= -\, P_{,i}$$

But,

$$\rho\, v^i{}_{,\mu} v^\mu = \rho\,(v^i{}_{,0} v^0 + v^i{}_{,j} v^j) \overset{(3.1\text{–}4)}{=} \rho\,(\frac{\partial v^i}{\partial (c\, t)}\; c + v^i{}_{,j} v^j) = \rho\,(\frac{\partial v^i}{\partial t} + v^i{}_{,j} v^j).$$

So, the $i$th component of equation (3.8) can be expressed as

$$\rho\,(\frac{\partial v^i}{\partial t} + v^i{}_{,j} v^j) = -\, P_{,i}\,,$$

which in 3-vector form is equation (3.9), $\rho\left(\frac{\partial}{\partial t} + \boldsymbol{v} \cdot \nabla\right)\boldsymbol{v} = -\nabla P$ :

$\frac{\partial \boldsymbol{v}}{\partial t}$ has $i$th component $\frac{\partial v^i}{\partial t}$.

$(\boldsymbol{v} \cdot \nabla)\boldsymbol{v} \overset{(3.1\text{–}14)}{=} v^j \frac{\partial \boldsymbol{v}}{\partial x^j}$ has $i$th component $v^j \frac{\partial v^i}{\partial x^j} = v^i{}_{,j} v^j$.

So, LHS of equation (3.9) has $i$th component $\rho\,(\frac{\partial v^i}{\partial t} + v^i{}_{,j} v^j)$. ✓

RHS has $i$th component $-\,\frac{\partial P}{\partial x^i} = -\, P_{,i}$ ✓ ■

The relativistic continuity equation (3.5) contains pressure while the classical version, equation (3.7), does not. That makes sense because energy rather than mass is conserved, and a fluid under pressure contributes to energy.

The relativistic equation of motion for a fluid (3.6) can be expressed as

$$(\rho + \frac{P}{c^2})\frac{d^2 x^\nu}{d\tau^2} = (\eta^{\mu\nu} - \frac{1}{c^2}u^\mu u^\nu)\, P_{,\mu} :$$

$$\frac{d^2 x^\nu}{d\tau^2} = \frac{d}{d\tau}\frac{dx^\nu}{d\tau} = (\frac{d}{d\tau}\delta^\nu_\mu)\frac{dx^\mu}{d\tau} = (\frac{d}{d\tau}\frac{\partial x^\nu}{\partial x^\mu})\frac{dx^\mu}{d\tau} = (\frac{\partial}{\partial x^\mu}\frac{dx^\nu}{d\tau})\frac{dx^\mu}{d\tau}$$

$$= \frac{\partial u^\nu}{\partial x^\mu}u^\mu = u^\nu{}_{,\mu}\, u^\mu$$

Thus, LHS of equation (3.6) can be written as $(\rho + \frac{P}{c^2})\frac{d^2 x^\nu}{d\tau^2}$ ✓

In this form, equation (3.6) looks more like an equation of motion, $F = ma$, pushing fluid particles off of geodesics (i.e., $\frac{d^2 x^\nu}{d\tau^2} = 0$) by the pressure gradient $P_{,\mu}$.

This finishes our derivation for flat spacetime. Using Principles 1 and 2 of Section 2.5, we can generalize these special relativity results to general relativity. This involves replacing $\eta^{\mu\nu}$ by $g^{\mu\nu}$ and replacing the partial derivative by the covariant derivative.

Definitions The **continuity equation** is

$$\boxed{(\rho\, u^\mu)_{;\mu} + \frac{P}{c^2}\, u^\mu{}_{;\mu} = 0} \;. \qquad (3.1\text{-}17)$$

The **equation of motion of a perfect fluid** is

$$\boxed{\left(\rho + \frac{P}{c^2}\right) u^\nu{}_{;\mu}\, u^\mu = \left(g^{\mu\nu} - \frac{1}{c^2}\, u^\mu\, u^\nu\right) P_{;\mu}} \;. \qquad (3.1\text{-}18)$$

The **stress tensor** for a perfect fluid is defined as

$$\boxed{T^{\mu\nu} \equiv \left(\rho + \frac{P}{c^2}\right) u^\mu\, u^\nu - P g^{\mu\nu}} \qquad (3.10)$$

and the vanishing of its divergence is expressed as

$$T^{\mu\nu}{}_{;\mu} = 0 \;. \qquad (3.11)$$

Recall Theorem 3.1.1, which easily generalizes to say that equation (3.11) holds iff both the continuity equation (3.1-17) and equation of motion equation (3.1-18) hold. Also, since energy and mass are equivalent in relativity, the relativistic continuity equation (3.1-17) now represents conservation of energy, not just conservation of fluid mass like its classical counterpart equation (3.7).

Associated energy-momentum tensors are

$$T_{\mu\nu} = (\rho + \frac{P}{c^2})\, u_\mu u_\nu - P g_{\mu\nu} \quad \text{and} \quad T_\mu^{\ \nu} = (\rho + \frac{P}{c^2})\, u_\mu u^\nu - P\delta_\mu^\nu : \qquad (3.1\text{-}19)$$

$$T_\mu^{\ \nu} = g_{\mu\sigma} T^{\sigma\nu} = (\rho + \frac{P}{c^2})\, g_{\mu\sigma} u^\sigma u^\nu - P\, g_{\mu\sigma} g^{\sigma\nu} = (\rho + \frac{P}{c^2})\, u_\mu u^\nu - P\delta_\mu^\nu \quad ✓$$

$$T_{\mu\nu} = g_{\nu\rho} T_\mu^{\ \rho} = (\rho + \frac{P}{c^2})\, g_{\nu\rho} u_\mu u^\rho - P g_{\nu\rho} g_\mu^\rho = (\rho + \frac{P}{c^2})\, u_\mu u_\nu - P g_{\nu\mu} \quad ✓$$

With suitable definition of $T^{\mu\nu}$, equation (3.11) is valid for all fluids and fields. The stress tensor will soon be shown to be the source of the gravitational field, but first we must discuss curvature. We should also point out that we define the stress to include all forces that exert pressure on the perfect fluid. That is, curved spacetime will result not just from gravitational force but from all forces.

## 3.2 The curvature tensor and related tensors

Unlike with partial derivatives, the order of differentiation matters with covariant derivatives. In this section, we develop the curvature tensor, $R$, from conditions that govern when the order matters, then we explore how $R$ relates to spacetime curvature, and we conclude with introduction of the associated Ricci and Einstein tensors. During this development, we only assume an $N$-dimensional manifold and, so, we use indices $a$, $b$, $c$, etc. instead of $\mu$, $\nu$, $\rho$. These results, however, apply to general relativity with $N = 4$.

The covariant derivative of a vector $\lambda_a$ is

$$\lambda_{a\,;\,b} \overset{(2.56)}{=} \partial_b \lambda_a - \Gamma^d_{a\,b} \lambda_d \tag{3.2-1}$$

and its second derivative is

$$\lambda_{a\,;\,b\,c} = \partial_c(\lambda_{a\,;b}) - \Gamma^e_{a\,c} \lambda_{e\,;\,b} - \Gamma^e_{b\,c} \lambda_{a\,;\,e} : \tag{3.2-2}$$

Let $\tau_{a\,b} \equiv \lambda_{a\,;\,b}$ . Then $\lambda_{a\,;\,b\,c} = \tau_{a\,b\,;\,c} \overset{(2.58)}{=} \partial_c \tau_{a\,b} - \Gamma^e_{a\,c} \tau_{e\,b} - \Gamma^e_{b\,c} \tau_{a\,e} = \text{RHS}$ ✓

Applying (3.2-1) to each term of (3.2-2) yields

$$\lambda_{a\,;\,b\,c} = \partial_c \partial_b \lambda_a - (\partial_c \Gamma^d_{a\,b})\, \lambda_d - \Gamma^d_{a\,b}\, \partial_c \lambda_d - \Gamma^e_{a\,c}\, [\partial_b \lambda_e - \Gamma^d_{e\,b} \lambda_d] - \Gamma^e_{b\,c}\, [\partial_e \lambda_a - \Gamma^d_{a\,e} \lambda_d] \,.$$

Interchanging $b$ and $c$ gives

$$\lambda_{a\,;\,c\,b} = \partial_b \partial_c \lambda_a - (\partial_b \Gamma^d_{a\,c})\, \lambda_d - \Gamma^d_{a\,c}\, \partial_b \lambda_d - \Gamma^e_{a\,b}\, [\partial_c \lambda_e - \Gamma^d_{e\,c} \lambda_d] - \Gamma^e_{c\,b}\, [\partial_e \lambda_a - \Gamma^d_{a\,e} \lambda_d] \,.$$

Subtracting gives

$$\lambda_{a\,;\,b\,c} - \lambda_{a\,;\,c\,b} = (\partial_b \Gamma^d_{a\,c} - \partial_c \Gamma^d_{a\,b} + \Gamma^e_{a\,c}\, \Gamma^d_{e\,b} - \Gamma^e_{a\,b}\, \Gamma^d_{e\,c})\, \lambda_d$$

since $\Gamma^e_{b\,c} = \Gamma^e_{c\,b}$. Define

$$\boxed{R^d{}_{a\,b\,c} \equiv \partial_b \Gamma^d_{a\,c} - \partial_c \Gamma^d_{a\,b} + \Gamma^e_{a\,c}\, \Gamma^d_{e\,b} - \Gamma^e_{a\,b}\, \Gamma^d_{e\,c}} \,. \tag{3.13}$$

Then $$\lambda_{a\,;\,b\,c} - \lambda_{a\,;\,c\,b} = R^d{}_{a\,b\,c}\, \lambda_d \tag{3.12}$$

The following theorem is an immediate consequence of equation (3.12).

Theorem 3.2.1 $\lambda_{a\,;\,b\,c} = \lambda_{a\,;\,c\,b}$ iff $R^d{}_{a\,b\,c} = 0$ for all $a$, $b$, $c$, $d$. That is, a necessary and sufficient condition for interchanging the order of covariant differentiation of covariant vector fields is that $R^d{}_{a\,b\,c} = 0$.

Theorem 3.2.2 $R^d{}_{abc}$ is a tensor.

Proof. $\lambda_{a;bc}$ is a tensor, so $R^d{}_{abc}\,\lambda_d = \lambda_{a;bc} - \lambda_{a;cb}$ is a tensor for arbitrary vectors $\lambda_d$. Thus, $R^d{}_{abc}$ satisfies the hypothesis of the Quotient Theorem 1.8.1, that contraction with an arbitrary vector generates a tensor. Thus, $R^d_{abc}$ is a tensor. ■

Since $R^d{}_{abc}$ is defined in terms of the metric tensor and its derivatives, we can expect it to be related to curvature. This is motivation for the following definition.

Definition **$R^d{}_{abc}$** is called the **curvature tensor** (or **Riemann-Christoffel tensor** or **Riemann tensor**).

The next theorem generalizes Theorem 3.2.1 from vectors to tensors.

Theorem 3.2.3 (Exercise 3.2.1) A necessary and sufficient condition for interchanging the order of covariant differentiation of tensor fields is that $R^d{}_{abc} = 0$.

In the solution to Exercise 3.2.1, we learned that the difference between the two 2nd order partial derivatives of a general tensor can be expressed

$$\boxed{\begin{aligned}&\tau^{a_1 \ldots a_r}{}_{b_1 \ldots b_s;cd} - \tau^{a_1 \ldots a_r}{}_{b_1 \ldots b_s;dc} \\ &= -\textstyle\sum_{k=1}^{r} R^{a_k}_{ecd}\,\tau^{a_1 \ldots a_{k-1}\,e\,a_{k+1} \ldots a_r}_{b_1 \ldots b_s} + \textstyle\sum_{k=1}^{s} R^{e}_{b_k\,cd}\,\tau^{a_1 \ldots a_r}{}_{b_1 \ldots b_{k-1}\,e\,b_{k+1} \ldots b_s}\end{aligned}} \tag{3.2-3}$$

Special cases that more easily reveal the pattern are

$$\lambda^a{}_{;bc} - \lambda^a{}_{;cb} = -R^a_{dbc}\,\lambda^d$$
$$\tau^{ab}{}_{;cd} - \tau^{ab}{}_{;dc} = -R^a_{ecd}\,\tau^{eb} - R^b_{ecd}\,\tau^{ae}$$
$$\tau^{ab}{}_{c;de} - \tau^{ab}{}_{c;ed} = -R^a_{fde}\,\tau^{fb}{}_c - R^b_{fde}\,\tau^{af}{}_c + R^f_{cde}\,\tau^{ab}{}_f$$
$$\lambda_{a;bc} - \lambda_{a;cb} = R^d_{abc}\,\lambda_d$$
$$\tau_{ab;cd} - \tau_{ab;dc} = R^e_{acd}\,\tau_{eb} + R^e_{bcd}\,\tau_{ae}$$

In flat Euclidean space using the Cartesian coordinate system, $\Gamma^a_{bc} = 0\ \forall\ a, b, c$. This implies that $R^a{}_{bcd} = 0$. A sphere, though embedded in flat Euclidean space, has non-zero $R^a{}_{bcd}$ in spherical coordinates. Based on such examples, we make the following definition.

Definition A **manifold is flat** if $R^a{}_{b\,c\,d} = 0$ at every point; otherwise, the **manifold is curved**.

By the Theorem 3.2.3, a manifold is flat iff we can interchange the order of the 1st and 2nd partial differentials of tensors.

Theorem 3.2.4 In any flat region of a manifold it is possible to introduce a Cartesian coordinate system (i.e., $g_{a\,b} = \eta_{a\,b}$).

At first glance, $R^a{}_{b\,c\,d}$ has $N^4$ components. However, $R^a{}_{b\,c\,d}$ possesses a number of symmetries that can be shown to cut the number down to $\frac{1}{12}N^2(N^2\text{-}1)$ . For spacetime, this is a reduction from 256 unknowns to 20.

Definition The equation

$$R^a{}_{b\,c\,d} + R^a{}_{c\,d\,b} + R^a{}_{d\,b\,c} = 0 \tag{3.14}$$

is known as the **cyclic identity**. (Exercise 3.2.2)

We saw in Section 1.8 that the metric tensors are used to create *associated tensors*.

Two **associated curvature tensors** are

$$R_{a\,b\,c\,d} \equiv g_{a\,e}\, R^e{}_{b\,c\,d} \quad \text{and} \quad R^{a\,b}{}_{c\,d} \equiv g^{b\,e}\, R^a{}_{e\,c\,d} \tag{3.2-4}$$

After extensive manipulation,

$$\begin{aligned} R_{a\,b\,c\,d} = {} & \tfrac{1}{2}\left[\partial_d\partial_a g_{b\,c} - \partial_d\partial_b g_{a\,c} + \partial_c\partial_b g_{a\,d} - \partial_c\partial_a g_{b\,d}\right] \\ & - g^{e\,f}\left[\Gamma_{e\,a\,c}\,\Gamma_{f\,b\,d} - \Gamma_{e\,a\,d}\,\Gamma_{f\,b\,c}\right] \end{aligned} \tag{3.15}$$

From equation (3.15) it is simple to check the following symmetry properties:

$$R_{a\,b\,c\,d} = - R_{b\,a\,c\,d} \tag{3.16}$$

$$R_{a\,b\,c\,d} = - R_{a\,b\,d\,c} \tag{3.17}$$

$$R_{a\,b\,c\,d} = R_{b\,a\,d\,c} \tag{3.2-5}$$

$$R_{a\,b\,c\,d} = R_{c\,d\,a\,b} \tag{3.18}$$

For example,

$$- R_{bacd} \overset{(3.15)}{=} -\frac{1}{2}\,[\partial_d\partial_b g_{ac} - \partial_d\partial_a g_{bc} + \partial_c\partial_a g_{bd} - \partial_c\partial_b g_{ac}]$$
$$+\ g^{ef}\,[\Gamma_{ebc}\,\Gamma_{fad} - \Gamma_{ebd}\,\Gamma_{fac}]$$
$$= \frac{1}{2}\,[-\partial_d\partial_b g_{ac} + \partial_d\partial_a g_{bc} - \partial_c\partial_a g_{bd} + \partial_c\partial_b g_{ac}]$$
$$-\ g^{ef}\,[-\,\Gamma_{fbc}\,\Gamma_{ead} + \Gamma_{fbd}\,\Gamma_{eac}]$$
$$\overset{(3.15)}{=} R_{abcd} \qquad \checkmark$$

Note: Take care that the curvature and other associated tensors may not necessarily mimic the behaviors (3.16 – 3.18 and 3.2-5):

To show this, we use the fact that $g^{ae}\,R_{ebcd} = R^a{}_{bcd}$ *raises b* and turns it into *a*. It is true that, for example, that

$$R^{ab}{}_{cd} \overset{\text{(raise)}}{=} g^{be} R^a{}_{ecd} \overset{\text{(raise)}}{=} g^{be} g^{af}\, R_{fecd} \overset{(3.2-5)}{=} g^{af}\, g^{be}\, R_{efdc}$$
$$\overset{\text{(raise)}}{=} g^{af} R^b{}_{fdc} \overset{\text{(raise)}}{=} R^{ba}{}_{dc} \qquad (3.2\text{-}6)$$

and

$$R^a{}_{bcd} \overset{\text{(raise)}}{=} g^{ae}\, R_{ebcd} \overset{(3.17)}{=} -\, R_{ebdc} \overset{\text{(raise)}}{=} -\, R^a{}_{bdc} \qquad (3.2\text{-}7)$$

but we run into possible difficulty with property (3.16) for the curvature tensor:

$$R^a{}_{bcd} \neq R_a{}^b{}_{cd} \overset{\text{(raise)}}{=} g^{be}\, R_{aecd} \overset{(3.16)}{=} -\ g^{be}\, R_{eacd} \overset{\text{(raise)}}{=} -\, R^b{}_{acd}\,.$$

Claim: $R^a{}_{acd} = 0$ : (3.19)

$$R^a{}_{acd} = g^{ba}\, R_{bacd} \ \text{ and } \ R^a{}_{acd} = g^{ab}\, R_{abcd} \overset{(3.16)}{=} -\, g^{ab}\, R_{bacd}$$
$$\Rightarrow \quad 2\, R^a{}_{acd} = g^{ba}\, R_{bacd} - g^{ab}\, R_{bacd} = (g^{ba} - g^{ab})\, R_{bacd} = 0 \qquad \checkmark$$

**Problem 3.1** In a 2-dimensional Riemannian manifold

$R_{1212} = R_{2121} = -R_{1221} = -R_{2112}$ and $R_{ABCD} = 0$ otherwise.

**Theorem 3.2.5** The covariant derivatives $R^a{}_{acd;e}$ satisfy the **Bianchi identity**

$$R^a{}_{bcd;e} + R^a{}_{bde;c} + R^a{}_{bec;d} = 0 : \qquad (3.20)$$

Proof.

First, $R^a{}_{bcd} \overset{(3.13)}{=} \partial_c\Gamma^a_{bd} - \partial_d\Gamma^a_{bc} + \Gamma^e_{bd}\,\Gamma^a_{ec} - \Gamma^e_{bc}\,\Gamma^a_{ed}\,.$

So, $R^a{}_{bcd;e} = \partial_e\partial_c\Gamma^a_{bd} - \partial_e\partial_d\Gamma^a_{bc}$

$$+\ (\partial_e\Gamma^e_{bd})\,\Gamma^a_{ec} + \Gamma^e_{bd}\,\partial_e\Gamma^a_{ec} - (\partial_e\Gamma^e_{bc})\,\Gamma^a_{ed} - \Gamma^e_{bc}\,\partial_e\Gamma^a_{ed}\,. \qquad (3.2\text{-}8)$$

At any point P we can construct a coordinate system with $(\Gamma^a_{b\,c})_\mathrm{P} = 0 \;\; \forall\, a, b, c$. However, this does not also mean that $(\partial_e \Gamma^a_{b\,c})_\mathrm{P} = 0$. Hence,

$$R^a{}_{b\,c\,d\,;\,e} \overset{(3.2\text{-}8)}{=} (\partial_e \partial_c \Gamma^a_{b\,d} - \partial_e \partial_d \Gamma^a_{b\,c})_{\;\mathrm{P}} \,.$$

Cyclically permuting $c$, $d$, and $e$ yields equation (3.20). Since P is arbitrary, equation (3.2) holds everywhere. Since a tensor equation that equals zero in one coordinate system equals zero in every coordinate system, the Bianchi identity holds in all coordinate systems. ■

Even though $R^a{}_{a\,c\,d} \overset{(3.19)}{=} 0$, in general $R^c{}_{a\,b\,c} \neq 0$. Thus, we make the following definition.

Definition The **Ricci tensor** is the contraction

$$\boxed{R_{a\,b} \equiv R^c{}_{a\,b\,c}} \,. \tag{3.21}$$

The Ricci tensor is symmetric in $a$ and $b$ (Exercise 3.2.4) :

$$R_{a\,b} = R_{b\,a} \tag{3.2-9}$$

$$0 \overset{(3.14)}{=} R^a{}_{b\,c\,a} + R^a{}_{c\,a\,b} + R^a{}_{a\,b\,c} \overset{(3.16)}{=} R^a{}_{b\,c\,a} + R^a{}_{c\,a\,b} \overset{(3.21)}{=} R_{b\,c} + R^a{}_{c\,a\,b}$$

$$\overset{(3.2\text{-}7)}{=} R_{b\,c} - R^a{}_{c\,b\,a} \overset{(3.21)}{=} R_{b\,c} - R_{c\,b} \qquad ✓$$

Thus, $R^a{}_b = g^{a\,c}\, R_{c\,b} = g^{a\,c}\, R_{b\,c} = R_b{}^a$, and we define the **associated Ricci tensor**

$$R^a_b \equiv R^a{}_b = R_b{}^a \tag{3.2-10}$$

Another **associated Ricci tensor** is

$$R^{a\,b} \equiv g^{b\,c}\, R^a_c. \tag{3.2-11}$$

The (1,1) Ricci tensor can also be contracted from the (2,2) curvature tensor.

$$R^a_b = R^{c\,a}{}_{b\,c} : \tag{3.2-12}$$

$$R^{c\,a}{}_{b\,c} \overset{(3.2\text{-}4)}{=} g^{a\,e}\, R^c{}_{e\,b\,c} \overset{(3.21)}{=} g^{a\,e}\, R_{e\,b} = R^a_b \;\; ✓$$

Definition The **curvature scalar** is $\boxed{R \equiv R^a_a = g^{a\,b}\, R_{b\,a}}$ (3.22)

Problem 3.1 (continued from above) In a Riemannian manifold, using spherical coordinates for a sphere of radius $a$,

$$R_{1212} = a^2 sin^2\theta \quad \text{and} \quad R = -\frac{2}{a^2}\,.$$

Definition The **Einstein tensor** is $\boxed{G_{ab} \equiv R_{ab} - \frac{1}{2} R\, g_{ab}}$ (3.23)

The Einstein tensor is important because it is the LHS of Einstein's field equations (sans cosmological constant). It is symmetric since $R$ and $g$ are. The **associated Einstein tensors** are

$$\boxed{G^a_b = R^a_b - \frac{1}{2} R\, \delta^a_b} \qquad (3.2\text{-}13)$$

and

$$\boxed{G^{ab} = G^a_c\, g^{cb} = R^{ab} - \frac{1}{2} R\, g^{ab}}: \qquad (3.2\text{-}14)$$

$$G^a_b = g^{ac}\, G_{cb} = g^{ac}\, (R_{cb} - \tfrac{1}{2}\, R\, g_{cb}) = R^a_b - \tfrac{1}{2}\, R\, \delta^a_b \qquad \checkmark$$
$$G^{ab} = G^a_c\, g^{cb} \;\; \checkmark \;\; \text{and} \;\; G^a_c\, g^{cb} = (R^a_c - \tfrac{1}{2}\, R\, \delta^a_c)\, g^{cb} = R^{ab} - \tfrac{1}{2}\, R\, g^{ab} \qquad \checkmark$$

Because $G_{ab}$ is symmetric, it has only one divergence, $G^{ab}{}_{;a}$. Divergence is the magnitude of "fluid" flow across the boundary of a region. Because the RHS of Einstein's field equations have zero divergence, the following theorem provides support for Einstein's invention of $G$ for the LHS.

Theorem 3.2.6 $G^{ab}{}_{;a} = 0$ .

Proof. Contracting $a$ with $d$ (i.e., subscript $d \mapsto a$) in the Bianchi identity (3.20) gives

$$R^a{}_{bca;e} + R^a{}_{bae;c} + R^a{}_{bec;a} = 0, \text{ or} \qquad (3.2\text{-}15)$$

We next generate expressions for the product of $g^{b\,e}$ with each of the three terms being differentiated in equation (3.2-15):

$$R^a{}_{b\,c\,a} \overset{(3.21)}{=} R_{b\,c} \quad \text{and} \quad R^a{}_{b\,a\,e} \overset{(3.2-7)}{=} -R^a{}_{b\,e\,a} \overset{(3.21)}{=} -R_{b\,e} \tag{3.2-16}$$

$$g^{b\,e}\,R^a{}_{b\,c\,a} \overset{(3.2-16)}{=} g^{b\,e}\,R_{b\,c} = R^e_c \tag{3.2-17}$$

$$g^{b\,e}\,R^a{}_{b\,a\,e} \overset{(3.2-16)}{=} -g^{e\,b}\,R_{b\,e} \overset{(3.22)}{=} -\mathrm{R} \tag{3.2-18}$$

$$g^{b\,e}\,R^a{}_{b\,e\,c} \overset{(3.2-4)}{=} R^{a\,e}{}_{e\,c} \overset{(3.2-6)}{=} R^{e\,a}{}_{c\,e} \overset{(3.2-12)}{=} R^a_c \tag{3.2-19}$$

In general, a variable like $g^{b\,e}$ cannot be moved in and out of a partial derivative like $R^a{}_{b\,c\,a\,;\,e}$ . However, we can assume we are using a coordinate system at a point P with $\Gamma^a_{b\,c} = 0 \;\; \forall\; a, b, c$. So, then $g^{b\,e} = \eta^{b\,e}$, a constant (i.e., +1, -1, or 0), and we are free, below, to move it in and out.

That is,

$$\begin{aligned} 0 &\overset{(3.2-15)}{=} g^{b\,e}\,(R^a{}_{b\,c\,a\,;\,e} + R^a{}_{b\,a\,e\,;\,c} + R^a{}_{b\,e\,c\,;\,a}) \overset{(3.2-17,\,3.2-18,\,3.2-19)}{=} R^e_{c\,;\,e} - R_{\,;\,c} + R^a_{c\,;\,a} \\ &= R^b_{c\,;\,b} - R_{\,;\,c} + R^b_{c\,;\,b} = 2R^b_{c\,;\,b} - (R\;\delta^b_c)_{\,;\,b} = (2R^b_c - R\;\delta^b_c)_{\,;\,b}\;, \end{aligned}$$

$$0 = (R^b_c - \tfrac{1}{2}R\;\delta^b_c)_{\,;\,b} \overset{(3.2-13)}{=} G^b_{c\,;\,b}\;, \tag{3.2-20}$$

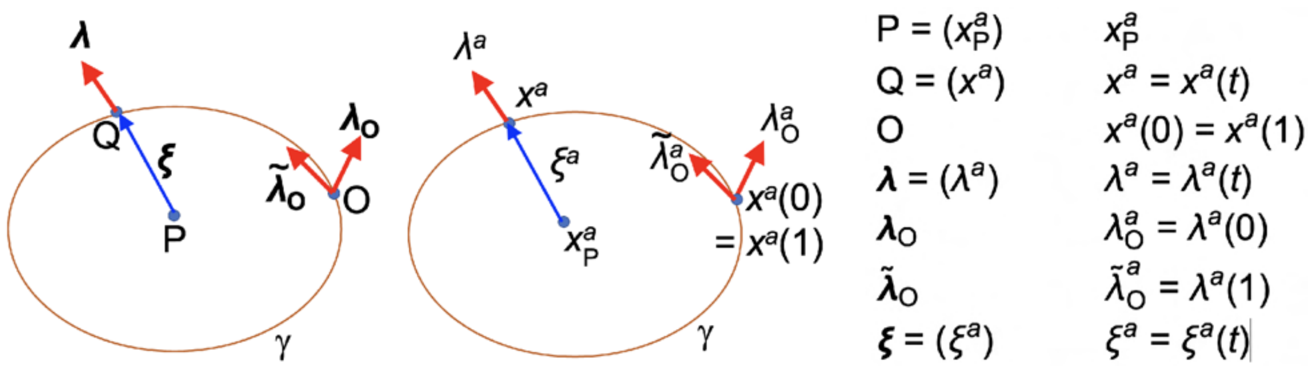
and

$$G^{b\,a}{}_{;\,b} = g^{\mathrm{ca}}\;G^b_{c\,;\,b} \overset{(3.2-20)}{=} 0. \tag{3.2-21}$$

Since P can be any point, divergence is zero everywhere. Finally, since divergence equaling zero is a tensor equation, it holds in every coordinate system. ■

## 3.3 Curvature and parallel transport

Parallel transport provides a method to compute the curvature components $R^a{}_{b\,c\,d}$. Specifically, in this section we develop an expression, equation (3.30), for $\Delta\lambda^a$ in terms of the curvature tensor $(R^a{}_{b\,c\,d})_{\mathrm{P}}$, where $\Delta\lambda^a$ is the change that results from parallelly transporting a vector $\lambda^a$ around a small closed loop $\gamma$ located about a point P in an $N$-manifold, $a = 1, \ldots, N$. This shows that $R^a{}_{b\,c\,d}$ is directly related to curvature effects and provides a justification for the name "curvature" tensor. We then illustrate how this equation can be used to find the components $R^a{}_{b\,c\,d}$.




**Points and Vectors** **Components** **Notation**

Let the curve $\gamma$ be parameterized by $t$: $0 \le t \le 1$, let $\lambda^a_{\mathrm{O}} = \lambda^a(0)$ be the initial vector, anchored at point O, and let $\tilde{\lambda}^a_{\mathrm{O}}$ be the vector after it is moved around the curve and back to point O by parallel transport. By definition (2.23) of parallel transport of $\lambda^a_{\mathrm{O}}$,

$$\frac{d\lambda^a}{dt} = -\Gamma^a_{b\,c}\,\lambda^b\,\frac{dx^c}{dt}\,. \tag{3.24}$$

By the Fundamental Theorem of Calculus, integrating equation (3.24) around the curve from point O to point Q gives

$$\lambda^a - \lambda^a_{\mathrm{O}} = \int_{\mathrm{O}}^{\mathrm{Q}} d\lambda^a \overset{(3.24)}{=} -\int_{\overset{\frown}{\mathrm{OQ}}} \Gamma^a_{b\,c}\,\lambda^b\,dx^c,$$

or

$$\lambda^a = \lambda^a_{\mathrm{O}} - \int_{\overset{\frown}{\mathrm{OQ}}} \Gamma^a_{b\,c}\,\lambda^b\,dx^c\,. \tag{3.25}$$

Let $\boldsymbol{\xi}$ be a small vector extending from P to Q. That is,

$$\xi^a \equiv x^a - x^a_{\mathrm{P}} \quad \text{or} \quad x^a = x^a_{\mathrm{P}} + \xi^a. \tag{3.3-1}$$

Then $dx^a \overset{(3.3-1)}{=} d\xi^a$ and, so,

$$\lambda^a \overset{(3.25)}{=} \lambda^a_O - \int_{\widehat{OQ}} \Gamma^a_{bc}\, \lambda^b\, d\xi^c \,. \tag{3.26}$$

Moving Q around the curve, from O around to O, gives

$$\tilde{\lambda}^a_O = \lambda^a_O - \oint \Gamma^a_{bc}\, \lambda^b\, d\xi^c \tag{3.3-2}$$

where $\tilde{\lambda}^a_O = \lambda^a(1)$, the vector at O after completing a circuit around $\gamma$.

Equation (3.3-2) provides a means for computing $\Delta\lambda^a = \tilde{\lambda}^a_O - \lambda^a_O$, but, unfortunately, $\lambda^b$ is also on RHS of the equation. However, for small loops, $\lambda^b \approx \lambda^b_O$. This is a zeroth order approximation. We plug in $\lambda^b_O$ for $\lambda^b$ on RHS of equation (3.26) to get a first order approximation; then plug in that result for $\lambda^b$ to get the next approximation, and so on. We will only need a second order approximation.

<u>Zeroth order approximation:</u> $\lambda^a = \lambda^a_O$

<u>First order approximation:</u>

$$\lambda^a = \lambda^a_O - \int_{\widehat{OQ}} \Gamma^a_{bc}\, \lambda^b_O\, d\xi^c = \lambda^a_O - \lambda^b_O \int_{\widehat{OQ}} \Gamma^a_{bc}\, d\xi^c$$

$$\Rightarrow \quad \lambda^b = \lambda^b_O - \lambda^d_O \int_{\widehat{OQ}} \Gamma^b_{de}\, d\xi^e \tag{3.3-3}$$

<u>Second order approximation:</u>

Replacing $\lambda^b$ in equation (3.3-2) by $\lambda^b$ from equation (3.3-3) yields:

$$\tilde{\lambda}^a_O = \lambda^a_O - \oint \Gamma^a_{bc} \left[ \lambda^b_O - \lambda^d_O \int_{\widehat{OQ}} \Gamma^b_{de}\, d\xi^e \right] d\xi^c$$

$$\tilde{\lambda}^a_O = \lambda^a_O - \lambda^b_O \oint \Gamma^a_{bc}\, d\xi^c + \lambda^d_O \oint \Gamma^a_{bc} \int_{\widehat{OQ}} \Gamma^b_{de}\, d\xi^e\, d\xi^c \tag{3.27}$$

$\Gamma^a_{bc}$ appears multiple time in equation (3.27), and we can express it as a Taylor series in order to generate 0th and 1st order approximations for it that we will then use to modify equation (3.27). To do this, recall from definition (2.13) that, though not obvious from the notation, $\Gamma^a_{b\,c}$ is a function of $x^a$, $x^b$, and $x^c$. By freezing $b$ and $c$, $\Gamma^a_{b\,c}$ becomes a function of $x^a$: $f(x^a) \equiv \Gamma^a_{b\,c}$ . Keep in mind that $\Gamma^a_{b\,c}$ and its partial derivatives are defined on the manifold, not along the curve $\gamma$.

The Taylor's series we develop will be expanded about the point P = $(x^a_{\mathrm{P}})$. Recall the Taylor series formula (2.2-32) for a function $f(x^a) = f\left(x^1, \ldots, x^N\right)$ of several variables:

$$f(x^a) = f(x^a_{\mathrm{P}}) + \frac{\partial}{\partial x^d} f(x^a_{\mathrm{P}})\,(x^d - x^d_{\mathrm{P}}) + \frac{1}{2!}\,\frac{\partial^2}{\partial x^d\,\partial x^e} f(x^a_{\mathrm{P}})\,(x^d - x^d_{\mathrm{P}})\,(x^e - x^e_{\mathrm{P}}) + \ldots\,.$$

Since $f(x^a) = \Gamma^a_{b\,c}$, then $\mathrm{f}(x^a_{\mathrm{P}}) = (\Gamma^a_{b\,c})_{\mathrm{P}}$. Recalling that $\xi^d \overset{(3.3\text{-}1)}{=} x^d - x^d_{\mathrm{P}}$, the Taylor series becomes

$$\Gamma^a_{b\,c} = (\Gamma^a_{b\,c})_{\mathrm{P}} + \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_{\mathrm{P}}\,\xi^d + \frac{1}{2!}\,\frac{\partial^2}{\partial x^d\,\partial x^e}(\Gamma^a_{b\,c})_{\mathrm{P}}\,\xi^d\,\xi^e + \ldots \qquad (3.3\text{-}4)$$

So, to the zeroth order, $\Gamma^a_{b\,c} = (\Gamma^a_{b\,c})_{\mathrm{P}}$ (3.3-5)

and to the first order in $\xi^d$, $\Gamma^a_{b\,c} = (\Gamma^a_{b\,c})_{\mathrm{P}} + \dfrac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_{\mathrm{P}}\,\xi^d$ . (3.3-6)

Examples, in general, of 1st order approximations are to approximate a circle by a line or a sphere by a plane. If we integrate a 1st order approximation over the circle or sphere, we ***improve***, achieving 2nd order approximations.

Before we substitute approximations for $\Gamma^a_{b\,c}$ in equation (3.27), we make a couple of observations. Applying the Fundamental Theorem of Calculus to $\boldsymbol{\xi}$ over the full loop yields

$$\oint d\xi^c = \xi^c(1) - \xi^c(0) = 0. \qquad (3.3\text{-}7)$$

Applying the Fundamental Theorem over just the arc $\overparen{\mathrm{OQ}}$ yields

$$\int_{\overparen{\mathrm{OQ}}} d\xi^e = \xi^e(\mathrm{t}) - \xi^e(0) = \xi^e - \xi^e(0)$$

$$\Rightarrow \quad \oint \int_{\overparen{\mathrm{OQ}}} d\xi^e\, d\xi^c = \oint [\xi^e - \xi^e(0]\, d\xi^c = [\,\oint \xi^e\, d\xi^c\,] - [\,\xi^e(0) \oint d\xi^c\,]$$

$$\overset{(3.3\text{-}7)}{=} \oint \xi^e\, d\xi^c \qquad (3.3\text{-}8)$$

Equation (3.27) is a 2nd order approximation. If we replace $\Gamma^a_{b\,c}$ in the 2nd term on RHS by a 1st order approximation, the 2nd term will remain part of a 2nd order approximation because the path integral “improves” the order of approximation. If we replace $\Gamma^a_{b\,c}$ in the 3rd term on RHS by a 0th order approximation, the 3rd term will remain part of a 2nd order approximation because of the double path integral. Specifically, for the second term,

$$\oint \Gamma^a_{b\,c}\, d\xi^c \;\mapsto\; (\Gamma^a_{b\,c})_P \oint d\xi^c + \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c \overset{(3.3\text{–}7)}{=} \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c$$

and for the third term,

$$\oint \Gamma^a_{b\,c} \int_{\widehat{0Q}} \Gamma^b_{d\,e}\, d\xi^e\, d\xi^c \;\mapsto\; (\Gamma^a_{b\,c})_P\, (\Gamma^b_{d\,e})_P \oint \int_{\widehat{0Q}} d\xi^e\, d\xi^c \overset{(3.3\text{–}8)}{=} (\Gamma^a_{b\,c})_P\, (\Gamma^b_{d\,e})_P \oint \xi^e\, d\xi^c.$$

Thus, the 2nd order approximation, equation (3.27), becomes

$$\tilde{\lambda}^a_O = \lambda^a_O - \lambda^b_O \frac{\partial}{\partial x^d}(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c + \lambda^d_O\, (\Gamma^a_{b\,c}\, \Gamma^b_{d\,e})_P \oint \xi^e\, d\xi^c,$$

and, to a 2nd order approximation,

$$\Delta\lambda^a = -\lambda^b_O\, [\, \partial_c(\Gamma^a_{b\,d})_P - (\Gamma^a_{e\,d}\, \Gamma^e_{b\,c})_P\, ] \oint \xi^c\, d\xi^d : \tag{3.28}$$

$$\begin{aligned} \Delta\lambda^a = \tilde{\lambda}^a_O - \lambda^a_O &= -\lambda^b_O\, \partial_d(\Gamma^a_{b\,c})_P \oint \xi^d\, d\xi^c + \lambda^d_O\, (\Gamma^a_{b\,c}\, \Gamma^b_{d\,e})_P \oint \xi^e\, d\xi^c \\ &= -\lambda^b_O\, \partial_c(\Gamma^a_{b\,d})_P \oint \xi^c\, d\xi^d + \lambda^b_O\, (\Gamma^a_{e\,d}\, \Gamma^e_{b\,c})_P \oint \xi^c\, d\xi^d \\ &= -\lambda^b_O\, [\, \partial_c(\Gamma^a_{b\,d})_P - (\Gamma^a_{e\,d}\, \Gamma^e_{b\,c})_P\, ] \oint \xi^c\, d\xi^d\,. \quad \checkmark \end{aligned}$$

Claim $\oint \xi^c\, d\xi^d = \frac{1}{2} \oint [\, \xi^c\, d\xi^d - \xi^d\, d\xi^c\, ]$ : (3.3-9)

(Exercise 3.3.1, part 1) By the Fundamental Theorem of Calculus,

$$\oint d(\xi^c \xi^d) = [\xi^c \xi^d]\,(1) - [\xi^c \xi^d]\,(0) = \xi^c(1)\, \xi^d(1) - \xi^c(0)\, \xi^d(0) = 0.$$

So,

$$0 = \oint d(\xi^c \xi^d) = \oint \xi^d\, d\xi^c + \oint \xi^c\, d\xi^d \tag{3.3-10}$$

$$\Rightarrow \qquad 2 \oint \xi^c\, d\xi^d = \oint [\, \xi^c\, d\xi^d - \xi^d\, d\xi^c\, ] \qquad \checkmark$$

Define a tensor $f^{c\,d}$ by

$$\boxed{f^{c\,d} \equiv \oint \xi^c\, d\xi^d \overset{(3.3\text{–}9)}{=} \frac{1}{2} \oint \left[\xi^c\, d\xi^d - \xi^d\, d\xi^c\right]}\,. \tag{3.29}$$

Observe that $f^{c\,d}$ is antisymmetric; i.e., $f^{c\,d} = - f^{d\,c}$ . (3.3-11)

Claim that the 2nd order approximation equation (3.28) can now be expressed as

$$\boxed{\Delta\lambda^a = -\tfrac{1}{2}\,(R^a{}_{b\,c\,d}\,)_P\,\lambda^b_O\,f^{c\,d}} \qquad (3.30)$$

where $f^{c\,d}$ is defined in equation (3.29):

$$- (\partial_c\Gamma^a_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c})_P\,f^{c\,d}$$

$$\overset{(3.3\text{-}11)}{=} -\frac{1}{2}(\partial_c\Gamma^a_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c}\,)_P\,f^{c\,d} + \frac{1}{2}(\partial_c\Gamma^a_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c}\,)_P\,f^{d\,c}$$

$$\overset{(c\leftrightarrow d)}{=} -\frac{1}{2}(\partial_c\Gamma^a_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c}\,)_P\,f^{c\,d} + \frac{1}{2}(\partial_d\Gamma^a_{b\,c} - \Gamma^a_{e\,c}\,\Gamma^e_{b\,d}\,)_P\,f^{c\,d}$$

$$- (\partial_c\Gamma^a_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c})_P\,f^{c\,d} = -\frac{1}{2}(\partial_c\Gamma^a_{b\,d} - \partial_d\Gamma^a_{b\,c} + \Gamma^a_{e\,c}\,\Gamma^e_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c}\,)_P\,f^{c\,d} \qquad (3.3\text{-}12)$$

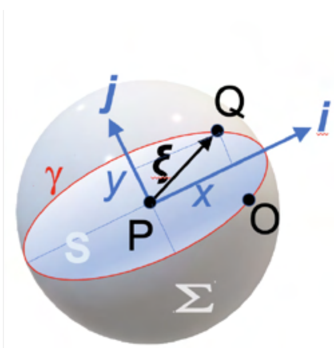
$$\Delta\lambda^a \overset{(3.28,\,3.29)}{=} -\lambda^b_O\,[\,\partial_c(\Gamma^a_{b\,d})_P - (\Gamma^a_{e\,d}\,\Gamma^e_{b\,c})_P\,]\,f^{c\,d}$$

$$\overset{(3.3\text{-}12)}{=} -\lambda^b_O\left(\tfrac{1}{2}\right)(\partial_c\Gamma^a_{b\,d} - \partial_d\Gamma^a_{b\,c} + \Gamma^a_{e\,c}\,\Gamma^e_{b\,d} - \Gamma^a_{e\,d}\,\Gamma^e_{b\,c}\,)_P\,f^{c\,d}$$

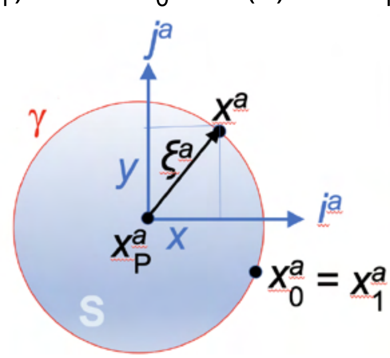
$$\overset{(3.13)}{=} -\frac{1}{2}\,(R^a{}_{b\,c\,d}\,)_P\,\lambda^b_O\,f^{c\,d} \qquad \checkmark$$

This concludes development of equation (3.30), which we now show can be exploited under certain conditions to solve for the components of the curvature tensor $R^a{}_{b\,c\,d}$ . We show this by developing equation (3.32), a geometric variation of equation (3.30), and then computing the components, Example 3.3.1 below.

Let P = $(x^a_P)$ be a point in an $N$-manifold M. Consider a small sphere Σ of radius $\rho$ about P, shown below. Form a loop $\gamma$ by intersecting a plane containing P with Σ. Let the loop be parameterized by $t$, $0 \le t \le 1$, so that points of $\gamma$ are expressed as $(x^a(t))$. Let the loop begin and end at point O = $(x^a_0)$ = $(x^a_1)$ where $x^a_0 = x^a(0)$ and $x^a_1 = x^a(1)$.




Points and Vectors




Components

Consider a pair of orthogonal unit vectors $\{\boldsymbol{i}, \boldsymbol{j}\}$ in the disk S enclosed by $\gamma$ and anchored at P . If the manifold M is flat, then $\{\boldsymbol{i}, \boldsymbol{j}\}$ constitutes a Cartesian coordinate system for S. If M is curved, then $\{\boldsymbol{i}, \boldsymbol{j}\}$ is approximately a Cartesian coordinate system for nearby points. A Cartesian coordinate system is a 1st order (linear) approximation. In particular, since $\gamma$ is the boundary of S, points of $\gamma$ can be expressed approximately to 1st order by

$$Q \approx P + x\, \boldsymbol{i} + y\, \boldsymbol{j}$$

where $x^2 + y^2 = \rho^2$. Note that $x$ and $y$ are functions of $t$ since Q is a function of $t$. Also $\boldsymbol{i}$ and $\boldsymbol{j}$ are fixed vectors. We express $\boldsymbol{i} = (\imath^a)$ and $\boldsymbol{j} = (j^a)$, with $\imath$ in a different font to clearly distinguish it from $j$). The equation above can now be expressed as

$$(x^a) \approx (x^a_{\mathrm{P}}) + x\,(\imath^a) + y\,(j^a).$$

In terms of components, the curve can be expressed

$$x^a \approx x^a_{\mathrm{P}} + x\, \imath^a + y\, j^a$$

where $\imath^a$ and $j^a$ are constants.

Let $\boldsymbol{\xi}$ be a (small) vector from P to a point Q = $(x^a)$ of the curve $\gamma$. In terms of components,

$$\xi^a = x^a - x^a_{\mathrm{P}} \approx x\, \imath^a + y\, j^a\ .$$

From equation (3.29),

$$f^{c\,d} = \frac{1}{2} \oint [\xi^c d\xi^d - \xi^d d\xi^c].$$

Claim $\xi^c d\xi^d - \xi^d d\xi^c = (x\, dy - y\, dx)\,(\imath^c j^d - \imath^d j^c)$ :

(Exercise 3.3.1, part 2)

$$\begin{aligned} \text{LHS} &= (x\, \imath^c + y\, j^c)\,(dx\, \imath^d + dy\, j^d) - (x\, \imath^d + y\, j^d)\,(dx\, \imath^c + dy\, j^c) \\ &= x\, dx\,(\imath^c \imath^d - \imath^d \imath^c) + y\, dy\,(j^c j^d - j^d j^c) \\ &\quad + x\, dy\,(\imath^c j^d - \imath^d j^c) - y\, dx\,(j^d \imath^c - j^c \imath^d) \\ &= (x\, dy - y\, dx)\,(\imath^c j^d - j^c \imath^d) \quad \checkmark \end{aligned}$$

Since the 1st order Cartesian approximation undergoes a path integral, it “improves” to a 2nd order approximation:

$$f^{c\,d} \overset{(3.29)}{=} \frac{1}{2} \oint [\,(x\, dy - y\, dx)\,(\imath^c j^d - \imath^d j^c)\,]$$

$$f^{c\,d} = \frac{1}{2}\,(\imath^c j^d - \imath^d j^c) \oint (x\, dy - y\, dx) \tag{3.31}$$

Now recall Green's Theorem (for Cartesian coordinates):

$$\oint_\gamma L\, dx + M\, dy = \int_S \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y}\right) dA$$

where $S$ is the disc enclosed by the loop $\gamma$, A is the area of S, and $L$ and $M$ are continuously differentiable functions of $x$ and $y$.

For $M = x$ and $L = 0$, $\oint_\gamma x\, dy = \int_S dA = A$

For $M = 0$ and $L = y$, $\oint_\gamma y\, dx = -\int_S dA = -A$

$$\Rightarrow \quad \frac{1}{2}\oint (x\, dy - y\, dx) = A \qquad (3.3\text{-}13)$$

On the (curved) manifold M, equation (3.3-13) is a 2nd order approximation (because it is a locally Euclidean, 1st order, approximation within a path integral). So, to the 2nd order we have

$$f^{c\,d} \overset{(3.31)}{=} A\,(\imath^c j^d - \imath^d j^c).$$

Since

$$\Delta\lambda^a \overset{(3.30)}{=} -\frac{1}{2}\,(R^a{}_{b\,c\,d})_P\, \lambda^b_O\, f^{c\,d}$$

to the 2nd order, then replacing $f^{c\,d}$ with a 2nd order approximation still yields a 2nd order approximation:

$$\frac{\Delta\lambda^a}{A} = -(R^a{}_{b\,c\,d})_P\, \lambda^b_O\, \imath^c j^d : \qquad (3.32)$$

$$\begin{aligned}
\Delta\lambda^a &= -\frac{1}{2}\,(R^a{}_{b\,c\,d})_P\, \lambda^b_O\, (\imath^c j^d - \imath^d j^c) \\
&= -\frac{1}{2}\,\lambda^b_O\, [\,(R^a{}_{b\,c\,d})_P\, \imath^c j^d - (R^a{}_{b\,c\,d})_P\, \imath^d j^c\,] \\
&= -\frac{1}{2}\,\lambda^b_O\, [\,(R^a{}_{b\,c\,d})_P\, \imath^c j^d - (R^a{}_{b\,d\,c})_P\, \imath^c j^d\,] \\
&\overset{(3.17)}{=} -\frac{1}{2}\,\lambda^b_O\, [\,(R^a{}_{b\,c\,d})_P\, \imath^c j^d - (R^a{}_{b\,c\,d})_P\, \imath^d j^c\,] \\
&= -\frac{1}{2}\,\lambda^b_O\, (R^a{}_{b\,c\,d})_P\, [\,\imath^c j^d + \imath^c j^d] \\
&= -(R^a{}_{b\,c\,d})_P\, \lambda^b_O\, \imath^c j^d \qquad ✓
\end{aligned}$$

In Problem 3.1 of Section 3.2 we showed how to compute the components of the curvature tensor $R_{a\,b\,c\,d}$ using the raw definitions, with the result that all non-zero components equal $\pm\, R_{1212}$. The next example shows another technique for calculating $R_{a\,b\,c\,d}$ that uses equation (3.32) along with judicious choices for $i^a$, $j^a$, and the parallel transport vector $\lambda_O^a$.

Example 3.3.1 Let P be the point with coordinates $(\theta_0, \phi_0)$ on a sphere of radius $a$, and let $\boldsymbol{\lambda}$ be a vector at P pointing south. We calculate $\Delta\boldsymbol{\lambda}$, the change in $\boldsymbol{\lambda}$ from parallelly transporting it around a small, closed loop about P.

Let the loop $\gamma$ be the sections of the latitude circles $\theta = \theta_0 \pm \epsilon$ and longitude circles $\phi = \phi_0 \pm \epsilon$ as shown in Figure 3.2.

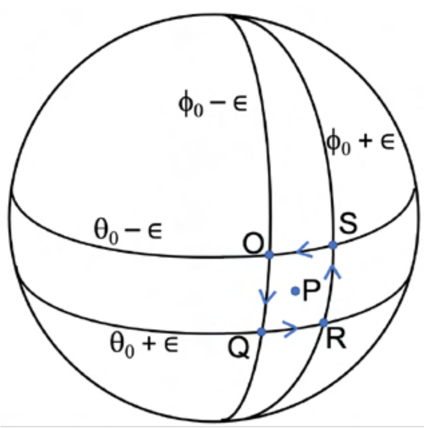



Figure 3.2

In Example 2.21 of Section 2.1 we learned that if a vector, located at a point $(\theta,\phi)$, points $\alpha$ radians east of south, then the vector has components

$$\boxed{\lambda^1 = \frac{\cos\alpha}{a} \text{ and } \lambda^2 = \frac{\sin\alpha}{a\sin\theta}}\,. \tag{2.25}$$

We also learned that if such a vector is parallely transported along a latitude arc from $(\theta,\phi)$ to $(\theta,\ \phi+\Delta\phi)$, then the final vector will have components

$$\boxed{\lambda^1 = \frac{\cos(\alpha-\omega)}{a} \text{ and } \lambda^2 = \frac{\sin(\alpha-\omega)}{a\sin\theta}}\,,\ \text{ where } \omega = \Delta\phi\,\cos\theta. \tag{2.2-10}$$

Of course, parallely transported vectors along longitude arcs don't change because longitudes are great circles, geodesics.

Start at point O with a unit vector $\boldsymbol{\lambda}_O$ pointing south and parallelly transport it from O to Q, R, S, and back to O. The coordinates of the initial south-pointing vector at O are

$$\lambda_O^1 = \frac{1}{a} \text{ and } \lambda_O^2 = 0,$$

$\boldsymbol{\lambda}$ arrives at, and then exits, point Q = $(\theta_0\text{-}\epsilon,\ \phi_0\text{-}\epsilon)$ still pointing south. So,

$$\lambda_Q^1 = \frac{1}{a} \text{ and } \lambda_Q^2 = 0.$$

The inputs to equation (2.2-10) at point R = $(\theta_0+\epsilon,\ \phi_0+\epsilon)$ are $\alpha = 0$, $\theta = \theta_0+\epsilon$, $\Delta\phi = 2\epsilon$, and $\omega = 2\epsilon \cos(\theta_0+\epsilon)$. So, the vector arrives at, and departs, R with

$$\lambda_R^1 = \frac{\cos[-2\,\epsilon\cos(\theta_0+\epsilon)]}{a} \text{ and } \lambda_R^2 = \frac{\sin[-2\,\epsilon\cos(\theta_0+\epsilon)]}{a\sin(\theta_0+\epsilon)}.$$

Next we transport along longitude arc $\overset{\frown}{RS}$, where the vector does not change:

$$\lambda_S^1 = \frac{\cos[-2\,\epsilon\cos(\theta_0+\epsilon)]}{a} \text{ and } \lambda_S^2 = \frac{\sin[-2\,\epsilon\cos(\theta_0+\epsilon)]}{a\sin(\theta_0-\epsilon)}.$$

We return to point O = $(\theta_0-\epsilon,\ \phi_0-\epsilon)$ where the inputs to equation (2.2-10) are $\alpha = -2\,\epsilon\cos(\theta_0+\epsilon)$, $\theta = \theta_0 - \epsilon$, $\Delta\phi = -2\epsilon$, and $\omega = -2\epsilon\cos(\theta_0 - \epsilon)$. So,

$$\begin{aligned}\alpha - \omega &= -2\,\epsilon\cos(\theta_0+\epsilon) + 2\epsilon\cos(\theta_0-\epsilon)\\ &= -2\epsilon\,(\cos\theta_0\cos\epsilon - \sin\theta_0\sin\epsilon) + 2\epsilon\,(\cos\theta_0\cos\epsilon + \sin\theta_0\sin\epsilon)\\ &= 4\epsilon\sin\theta_0\sin\epsilon,\end{aligned}$$

and the final vector is

$$\tilde{\lambda}_O^1 = \frac{\cos(4\,\epsilon\sin\theta_0\sin\epsilon)}{a} \text{ and } \tilde{\lambda}_O^2 = \frac{\sin(4\,\epsilon\sin\theta_0\sin\epsilon)}{a\sin(\theta_0-\epsilon)}.$$

To use equation (3.32) to approximate the curvature tensor, we need a 2nd order approximation for $\Delta\boldsymbol{\lambda}_P$ as well as for the area A enclosed by the 4 arcs, and we will need to choose a pair of orthogonal unit vectors $\{\boldsymbol{i}, \boldsymbol{j}\}$.

Recall that the 2nd order Taylor series approximations for sine and cosine are

$$\sin\theta = \theta \quad\text{and}\quad \cos\theta = 1 - \tfrac{1}{2}\theta^2.$$

From this we get additional 2nd order approximations:

$sin(\theta_0 - \epsilon) = \theta_0 - \epsilon = \theta_0, \qquad 4\epsilon \, sin\ \theta_0 \ sin\ \epsilon = 4\epsilon^2 \theta_0,$

$sin(4\epsilon \, sin\theta_0 \ sin\ \epsilon) = sin(4\epsilon^2 \theta_0) = 4\epsilon^2 \theta_0,$

$cos(4\epsilon \, sin\theta_0 \ sin\ \epsilon) = cos(4\epsilon^2 \theta_0) = 1 + \frac{1}{2}\,(16\epsilon^4)\ \theta_0^2 = 1.$

Thus, to the 2nd order in $\epsilon$, (Exercise 3.3.3)

$$\Delta\lambda^1 \approx \Delta\lambda^1_O = \tilde{\lambda}^1_O - \lambda^1_O = \frac{cos\,(4\,\epsilon\, sin\,\theta_0\, sin\,\epsilon\,)}{a} - \frac{1}{a} = \frac{1}{a} - \frac{1}{a} = 0,$$

$$\Delta\lambda^2 \approx \Delta\lambda^2_O = \tilde{\lambda}^2_O - \lambda^2_O = \frac{sin\,(4\,\epsilon\, sin\,\theta_0\, sin\,\epsilon\,)}{a\, sin\,(\theta_0 - \epsilon)} - 0 = \frac{4\,\epsilon^2\,\theta_0}{a\,\theta_0} = \frac{4\,\epsilon^2}{a};$$

i.e.,

$$\Delta\lambda^a = \frac{4\,\epsilon^2}{a}\,\delta^a_2. \qquad (3.3\text{-}14)$$

We next generate a 2nd order approximation for the area A enclosed by the 4 arcs. While the equator has radius $a$, a circle at latitude $\hat{\theta} \neq \frac{\pi}{2}$ has a smaller radius, namely $r = a\ sin\hat{\theta}$. For latitude arc $\overparen{QR}$, $r = a\ sin(\theta_0+\epsilon)$ and the angle between the center of the sphere and points Q and R is $\theta \approx 2\epsilon$. Then $s \equiv r\theta = 2\epsilon\ a\ sin(\theta_0+\epsilon)$. For latitude arc $\overparen{OS}$, $s = 2\epsilon\ a\ sin(\theta_0-\epsilon)$. So the lengths of the top and bottom of the "square" enclosed by the four arcs is $s = 2\epsilon\ a\ sin\ \theta_0$ to a 1st order approximation.

Every longitude circle has radius $a$, so the sides of the "square" have length $s = 2\epsilon a$, also to a 1st order approximation. Consequently, the area is

$$A = (2\epsilon\ a\ sin\ \theta_0)\ (2\epsilon a) = 4\epsilon^2 a^2 sin\ \theta_0, \qquad (3.3\text{-}15)$$

a 2nd order approximation in $\epsilon$.

Finally, a simple choice for ***i*** and ***j*** are the unit vectors in the tangent directions $\boldsymbol{e}_\theta$ and $\boldsymbol{e}_\phi$. For spherical coordinates, from Example 1.1.4,

$$\boldsymbol{e}_\theta = a\ cos\theta\ cos\phi\ \boldsymbol{i} + a\ cos\theta\ sin\phi\ \boldsymbol{j} - a\ sin\theta\ \boldsymbol{k} \qquad (3.3\text{-}16)$$

$$\boldsymbol{e}_\phi = -\ a\ sin\theta\ sin\phi\ \boldsymbol{i} + a\ sin\theta\ cos\phi\ \boldsymbol{j}\ . \qquad (3.3\text{-}17)$$

So, $|\boldsymbol{e}_\theta| = a$ and $|\boldsymbol{e}_\phi| = a\ sin\ \theta_0$. We select orthonormal unit vectors

$$i^1 = \frac{1}{a}, \quad i^2 = 0, \quad j^1 = 0, \quad \text{and} \quad j^2 = \frac{1}{a\, sin\,\theta_0},$$

or

$$i^a = \frac{1}{a}\,\delta_1^a \quad \text{and} \quad j^a = \frac{1}{a\,sin\,\theta_0}\,\delta_2^a. \tag{3.3-18}$$

We now have all the pieces we need to use equation (3.32) to find $(R^a{}_{b\,c\,d})_\mathrm{P}$.

$$\frac{\Delta\,\lambda^a}{A} = \overset{(3.3-14,\,3.3-15)}{=} \frac{4\,\epsilon^2}{a}\,\delta_2^a\,\frac{1}{4\,\epsilon^2\,a^2\,sin\,\theta_0} = \delta_2^a\,\frac{1}{a^3\,sin\,\theta_0}. \tag{3.3-19}$$

The vector, $\boldsymbol{\lambda}$, we have been transporting started at point O and points south. So, $\lambda_O^b = \frac{1}{a}\,\delta_1^b$, and

$$\begin{aligned} -\,(R^a{}_{b\,c\,d})_\mathrm{P}\,\lambda_O^b\,i^c j^d &\overset{(3.3-18)}{=} -\,(R^a{}_{b\,c\,d})_\mathrm{P}\,\frac{1}{a}\,\delta_1^b\,\frac{1}{a}\,\delta_1^c\,\frac{1}{a\,sin\,\theta_0}\,\delta_2^d \\ &= -\,(R^a{}_{112})_\mathrm{P}\,\frac{1}{a^3\,sin\,\theta_0} \end{aligned} \tag{3.3-20}$$

$$\delta_2^a\,\frac{1}{a^3\,sin\,\theta_0} \overset{(3.3-19)}{=} \frac{\Delta\,\lambda^a}{A} \overset{(3.32)}{=} -\,(R^a{}_{b\,c\,d})_\mathrm{P}\,\lambda_O^b\,i^c j^d \overset{(3.3-20)}{=} -\,(R^a{}_{112})_\mathrm{P}\,\frac{1}{a^3\,sin\,\theta_0}$$

$$\Rightarrow \quad (R^a{}_{112})_\mathrm{P} = -\,\delta_2^a \tag{3.3-21}$$

$$\Leftrightarrow \quad (R^1{}_{112})_\mathrm{P} = 0 \quad \text{and} \quad (R^2{}_{112})_\mathrm{P} = -1 \tag{3.3-22}$$

This enables us to compute all of the Curvature tensor components:

From 3.3-16 and 3.3-17:

$$g_{11} = \mathbf{e}_{\theta_0} \bullet \mathbf{e}_{\theta_0} = a^2 \tag{3.3-23}$$

$$g_{12} = g_{21} = \mathbf{e}_{\theta_0} \bullet \mathbf{e}_{\phi_0} = 0 \tag{3.3-24}$$

$$g_{22} = \mathbf{e}_{\phi_0} \bullet \mathbf{e}_{\phi_0} = a^2\,sin^2\,\theta_0 \tag{3.3-25}$$

$$(R_{a\,112})_\mathrm{P} = g_{a\,e}\,(R^e{}_{112})_\mathrm{P} \overset{(3.3-21)}{=} -\,g_{a\,e}\,\delta_2^e = -\,g_{a\,2}$$

For $a = 2$, $(R_{1212})_\mathrm{P} = -\,(R_{2112})_\mathrm{P} = g_{22} \overset{(3.3-25)}{=} a^2\,sin^2\,\theta_0$.

By Problem 3.1, $(R_{1212})_\mathrm{P} = (R_{2121})_\mathrm{P} = -(R_{2112})_\mathrm{P} = -(R_{1221})_\mathrm{P} = g_{22}$, and the remaining components are zero.

For $a = 1$, $(R_{1112})_\mathrm{P} = g_{21} = 0$, which is in agreement with the remaining components being zero. ■

The following observations are not found in the book. $R_{1212} = a^2\,sin^2\,\theta_0$ is interesting because in classical physics for a sphere of radius $a$, the (scalar) curvature is $\kappa = \frac{1}{a}$. So, $R_{1212}$ appears to be related. Pursuing this further:

$$\mathbf{e}^{\theta} = \frac{\cos\theta\ \cos\phi}{a}\ \boldsymbol{i} + \frac{\cos\theta\ \sin\phi}{a}\ \boldsymbol{j} - \frac{\sin\theta}{a}\ \boldsymbol{k}$$

$$\mathbf{e}^{\phi} = -\frac{\sin\phi}{a\sin\theta}\ \boldsymbol{i} + \frac{\cos\phi}{a\sin\theta}\ \boldsymbol{j}$$

$$g^{11} = \mathbf{e}^{\theta_0} \bullet \mathbf{e}^{\theta_0} = \frac{1}{a^2}$$

$$g^{12} = g^{21} = \mathbf{e}^{\theta_0} \bullet \mathbf{e}^{\phi_0} = 0$$

$$g^{22} = \mathbf{e}^{\phi_0} \bullet \mathbf{e}^{\phi_0} = \frac{1}{a^2 \sin^2\theta_0}$$

$$R_{a\,b} \overset{(3.21)}{=} R^{c}_{\ a\,b\,c} = g^{e\,c}\, R_{e\,a\,b\,c}$$

$$\Rightarrow \quad R_{11} = g^{e\,c} R_{e11c}$$

The only non-zero value of $R_{e11c}$ occurs when $e = c = 2$:

$$R_{11} = g^{22} R_{2112} = \frac{1}{a^2 \sin^2\theta_0}\ (-\ a^2 \sin^2\theta_0) = -\ 1$$

The only non-zero value of $R_{e12c}$ occurs when $e = 2$ and $c = 1$:

$$R_{12} = g^{e\,c} R_{e12c} = g^{21} R_{2121} = (0)\ (a^2 \sin^2\theta_0) = 0 \ \text{ and } \ R_{21} = 0$$

The only non-zero value of $R_{e22c}$ occurs when $e = c = 1$:

$$R_{22} = g^{e\,c} R_{e22c} = g^{11} R_{1221} = \frac{1}{a^2}\ (-\ a^2 \sin^2\theta_0) = -\ \sin^2\theta_0$$

$$R \overset{(3.22)}{=} g^{a\,b}\, R_{a\,b} = g^{11}\, R_{11} + g^{22}\, R_{22} = \frac{1}{a^2}\ (-1) + \frac{1}{a^2 \sin^2\theta_0}\ (-\ \sin^2\theta_0) = -\ \frac{2}{a^2}.$$

So, though they may be related, $\kappa$ and $R$ are not the same. Moreover, it is useful to have an example where we have values of all the curvature tensors and associated tensors. To this end we include the following tensor values:

$R^{a\,b} = g^{b\,c}\, R^{a}_{c} = g^{a\,d} g^{b\,c}\, R_{d\,c}$ :

$R_{12} = R_{21} = 0$, so we need only consider terms with $R_{11}$ and $R_{22}$.

$$R^{11} = g^{1\,d} g^{2\,c}\, R_{d\,c} = g^{11} g^{21}\, R_{11} + g^{12} g^{22}\, R_{22} = \frac{1}{a^4}(-1) + 0 = -\frac{1}{a^4}$$

$$R^{12} = g^{1\,d} g^{1\,c}\, R_{d\,c} = g^{11} g^{11}\, R_{11} + g^{12} g^{12}\, R_{22} = 0 \ \text{ and } \ R^{21} = 0$$

$$R^{22} = g^{2\,d} g^{2\,c}\, R_{d\,c} = g^{21} g^{21}\, R_{11} + g^{22} g^{22}\, R_{22} = 0 + \frac{1}{a^4 \sin^4\theta_0}(-\sin^2\theta_0)$$

$$= -\frac{1}{a^4 \sin^2\theta_0}$$

$R^a_b = g^{a\,c} R_{c\,b}$ :

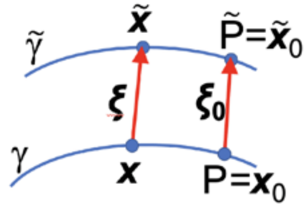
Again, we only need to consider terms with $R_{11}$ and $R_{22}$.

$R^1_1 = g^{11} R_{11} = \frac{1}{a^2}\ (-1) = -\ \frac{1}{a^2}$

$R^1_2 = g^{12} R_{22} = 0$ and $R^2_1 = g^{21} R_{11} = 0$

$R^2_2 = g^{22} R_{22} = \frac{1}{a^2\, sin^2\, \theta_0}\ (-\ sin^2\, \theta_0) = -\ \frac{1}{a^2}$

## 3.4 Geodesic Deviation




Another layer of support for Einstein's field equations come from the equation of geodesic deviation, 3.35. We derive the equation and then show that it reduces to its Euclidean counterpart.

Let P be a point in a manifold M. Since M is locally Euclidean, there is an open neighborhood U about P that is homeomorphic to Euclidean space. Thus, P has a Euclidean coordinate system ($x^a$) defined on U. Let $\tilde{P}$ be a different point in U. Then $\tilde{P}$ also has a Euclidean coordinate system, ($\tilde{x}^a$), defined on U.

Let $\gamma$ be a geodesic curve through P that is contained in U, and let $\tilde{\gamma}$ be a geodesic curve through $\tilde{P}$ that is also contained in U. Let both curves be parameterized by $u$ such that P = $\boldsymbol{x}_0(0)$ and $\tilde{P}$ = $\tilde{\boldsymbol{x}}_0(0)$. Denote $\boldsymbol{x}_0 = \boldsymbol{x}_0(0)$ and $\tilde{\boldsymbol{x}}_0 = \tilde{\boldsymbol{x}}_0(0)$. Typical points on $\gamma$ and $\tilde{\gamma}$ are denoted $\boldsymbol{x} = \boldsymbol{x}(u)$ and $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}(u)$, respectively. One of the curves may encompass a larger range for $u$ than the other, so we restrict ourselves to a range of $u$ that is valid for both curves.

Define a "vector" $\boldsymbol{\xi} = \boldsymbol{\xi}(u) = \tilde{\boldsymbol{x}} - \boldsymbol{x}$ and denote $\boldsymbol{\xi}_0 = \tilde{\boldsymbol{x}}_0 - \boldsymbol{x}_0$. ($\boldsymbol{\xi}$ is not actually a vector because it does not lie in the tangent plane at either $\boldsymbol{x}$ or $\tilde{\boldsymbol{x}}$.) Recall that curve $\gamma$ is null if $g_{ab}\,\dot{x}^a\,\dot{x}^b = 0$ at any point, where "dot" refers to differentiation by $u$. If neither $\gamma$ nor $\tilde{\gamma}$ is null, then by Exercise 2.1.3 we can let $u$ be an affine parameter $u = As + B$, where $s$ is arc length. If $\gamma$ is null then for some $x^a$, then $ds \overset{(1.83)}{=} \sqrt{\left|\, g_{ab}\,\dot{x}^a\,\dot{x}^b \,\right|} = 0$
$\Rightarrow$ $s$ is a constant, not suitable as a parameter on $\gamma$.

Since $\gamma$ and $\tilde{\gamma}$ are geodesics, they satisfy the geodesic equation (2.72):

$$\frac{d^2\,\tilde{x}^a}{du^2} + \tilde{\Gamma}^a_{bc}\,\frac{d\,\tilde{x}^b}{du}\,\frac{d\,\tilde{x}^c}{du} = 0 \quad \text{or} \quad \ddot{\tilde{x}}^a + \tilde{\Gamma}^a_{bc}\,\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c = 0 \tag{3.33}$$

$$\frac{d^2\,x^a}{du^2} + \Gamma^a_{bc}\,\frac{dx^b}{du}\,\frac{dx^c}{du} = 0 \quad \text{or} \quad \ddot{x}^a + \Gamma^a_{bc}\,\dot{x}^b\,\dot{x}^c = 0 \tag{3.34}$$

where

$$\tilde{\Gamma}^a_{bc} \equiv (\Gamma^a_{bc})_{\tilde{x}} \quad \text{and} \quad \Gamma^a_{bc} \equiv (\Gamma^a_{bc})_{x}.$$

We next derive equations labeled (3.4-2) – (3.4-5), and (3.35). These derivations constitute Exercise 3.4.1 in the book.

For $\xi^a = \tilde{x}^a - x^a$, it was shown in Section 3.3 that to the 1st order in $\xi^a$,

$$\tilde{\Gamma}^a_{bc} \overset{(3.3\text{–}6)}{=} \Gamma^a_{bc} + \left(\frac{\partial}{\partial x^d}\Gamma^a_{bc}\right)\xi^d = \Gamma^a_{bc} + \Gamma^a_{bc,d}\,\xi^d.$$

So, equation (3.33) can be rewritten as

$$\ddot{\tilde{x}}^a + (\Gamma^a_{bc} + \Gamma^a_{bc,d}\,\xi^d)\,\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c = 0 \tag{3.4-1}$$

Also,

$$\tilde{x}^a = x^a + \xi^a, \quad \dot{\tilde{x}}^a = \dot{x}^a + \dot{\xi}^a, \text{ and } \quad \ddot{\tilde{x}}^a = \ddot{x}^a + \ddot{\xi}^a.$$

Subtracting (3.34) from (3.4-1) yields

$$\ddot{\tilde{x}}^a - \ddot{x}^a + \Gamma^a_{bc}\,(\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c - \dot{x}^b\,\dot{x}^c) + \Gamma^a_{bc,d}\,\xi^d\,\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c = 0. \tag{3.4-2}$$

Claim: equation (3.4-2) can be expressed to the 1st order in $\xi^a$ as

$$\ddot{\xi}^a + \Gamma^a_{bc,d}\,\dot{x}^b\,\dot{x}^c\,\xi^d + \Gamma^a_{bc}\,(\dot{x}^b\dot{\xi}^c + \dot{\xi}^b\dot{x}^c) = 0 : \tag{3.4-3}$$

$$\ddot{\tilde{x}}^a - \ddot{x}^a = \ddot{\xi}^a$$

$$\Gamma^a_{bc}\,(\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c - \dot{x}^b\,\dot{x}^c) = \Gamma^a_{bc}\,[\,(\dot{x}^b + \dot{\xi}^b)\,(\dot{x}^c + \dot{\xi}^c) - \dot{x}^b\,\dot{x}^c] = \Gamma^a_{bc}\,(\dot{x}^b\dot{\xi}^c + \dot{\xi}^b\,\dot{x}^c)$$

$$\begin{aligned}\Gamma^a_{bc,d}\,\xi^d\,\dot{\tilde{x}}^b\,\dot{\tilde{x}}^c &= \Gamma^a_{bc,d}\,\xi^d\,(\dot{x}^b + \dot{\xi}^b)\,(\dot{x}^c + \dot{\xi}^c)\\ &= \Gamma^a_{bc,d}\,\xi^d\,(\dot{x}^b\,\dot{x}^c + \dot{x}^b\dot{\xi}^c + \dot{\xi}^b\dot{x}^c + \dot{\xi}^b\dot{\xi}^c)\\ &= \Gamma^a_{bc,d}\,\dot{x}^b\,\dot{x}^c\,\xi^d\end{aligned}$$

to the 1st order since $\xi^d\,(\dot{x}^b\dot{\xi}^c + \dot{\xi}^b\dot{x}^c)$ is 2nd order in $\xi^a$. I think this is 2nd order because $\xi^d\xi^c$ is 2nd order, and so $\xi^d\dot{\xi}^c$ is at least 2nd order. ✓

Claim: equation (3.4-3) can be rewritten as

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{b\,c}\,\xi^b\dot{x}^c\right) - \Gamma^a_{b\,c,\,d}\,\xi^b\dot{x}^c\dot{x}^d \;-\; \Gamma^a_{b\,c}\,\xi^b\ddot{x}^c$$
$$+\,\Gamma^a_{b\,c,\,d}\,\dot{x}^b\dot{x}^c\xi^d \;+\,\Gamma^a_{b\,c}\,\dot{x}^b\dot{\xi}^c = 0: \qquad (3.4\text{-}4)$$

$$\frac{d}{du}\Gamma^a_{b\,c} = \frac{\partial}{\partial x^d}\Gamma^a_{b\,c}\,\frac{dx^d}{du} = \Gamma^a_{b\,c,\,d}\,\dot{x}^d$$
$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{b\,c}\,\xi^b\dot{x}^c\right) = \ddot{\xi}^a + \frac{d}{du}\Gamma^a_{b\,c}\,\xi^b\dot{x}^c \;+\Gamma^a_{b\,c}\,\dot{\xi}^b\dot{x}^c +\; \Gamma^a_{b\,c}\,\xi^b\ddot{x}^c$$
$$= \ddot{\xi}^a + \Gamma^a_{b\,c,\,d}\,\xi^b\dot{x}^c\dot{x}^d + \Gamma^a_{b\,c}\,\dot{\xi}^b\dot{x}^c + \;\Gamma^a_{b\,c}\,\xi^b\ddot{x}^c$$

Plugging the expression above into equation (3.4-4) yields equation (3.4-3):

$$0 = \ddot{\xi}^a + \;\Gamma^a_{b\,c,\,d}\,\xi^b\dot{x}^c\dot{x}^d + \Gamma^a_{b\,c}\,\dot{\xi}^b\dot{x}^c + \;\Gamma^a_{b\,c}\,\xi^b\ddot{x}^c$$
$$-\;\Gamma^a_{b\,c,\,d}\,\xi^b\dot{x}^c\dot{x}^d - \;\Gamma^a_{b\,c}\,\xi^b\ddot{x}^c + \Gamma^a_{b\,c,\,d}\,\dot{x}^b\dot{x}^c\xi^d + \Gamma^a_{b\,c}\,\dot{x}^b\dot{\xi}^c$$
$$= \ddot{\xi}^a + \;\Gamma^a_{b\,c,\,d}\,\dot{x}^b\dot{x}^c\xi^d \;+\;\Gamma^a_{b\,c}\left(\dot{x}^b\dot{\xi}^c + \dot{\xi}^b\dot{x}^c\right) \qquad ✓$$

Claim: equation (3.4-4) can be written

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{b\,c}\,\xi^b\dot{x}^c\right) + \Gamma^a_{d\,e}\left(\dot{\xi}^d + \Gamma^d_{b\,c}\,\xi^b\dot{x}^c\right)\dot{x}^e - \Gamma^a_{d\,e}\,\Gamma^d_{b\,c}\,\xi^b\dot{x}^c\dot{x}^e$$
$$-\;\Gamma^a_{b\,c,\,d}\,\xi^b\dot{x}^c\dot{x}^d + \Gamma^a_{b\,c}\,\xi^b\,\Gamma^c_{d\,e}\,\dot{x}^d\dot{x}^e + \Gamma^a_{b\,c,\,d}\,\dot{x}^b\dot{x}^c\xi^d = 0: \qquad (3.4\text{-}5)$$

Equation (3.34) can be expressed as $\ddot{x}^c = -\,\Gamma^c_{d\,e}\,\dot{x}^d\dot{x}^e$ . So, substituting the two expressions below into equation (3.4-4) yields equation (3.4-5):

$$-\,\Gamma^a_{b\,c}\,\xi^b\ddot{x}^c \overset{(3.34)}{=} \Gamma^a_{b\,c}\,\xi^b\,\Gamma^c_{d\,e}\,\dot{x}^d\dot{x}^e \qquad ✓$$
$$\Gamma^a_{b\,c}\,\dot{x}^b\dot{\xi}^c \;=\; \Gamma^a_{e\,d}\,\dot{x}^e\dot{\xi}^d = \Gamma^a_{d\,e}\dot{\xi}^d\dot{x}^e$$
$$= \Gamma^a_{d\,e}\dot{\xi}^d\dot{x}^e + \Gamma^a_{d\,e}\,\Gamma^d_{b\,c}\,\xi^b\dot{x}^c\dot{x}^e - \Gamma^a_{d\,e}\,\Gamma^d_{b\,c}\,\xi^b\dot{x}^c\dot{x}^e$$
$$= \Gamma^a_{d\,e}\left(\dot{\xi}^d + \Gamma^d_{b\,c}\,\xi^b\dot{x}^c\right)\dot{x}^e - \Gamma^a_{d\,e}\,\Gamma^d_{b\,c}\,\xi^b\dot{x}^c\dot{x}^e \qquad ✓$$

From $\dfrac{D\xi^a}{du} \overset{(2.45)}{=} \dot{\xi}^a + \Gamma^a_{b\,c}\,\xi^b\dot{x}^c$ we see that the first two terms of equation (3.4-5) can be expressed as

$$\frac{d}{du}\left(\dot{\xi}^a + \Gamma^a_{b\,c}\,\xi^b\dot{x}^c\right) + \Gamma^a_{d\,e}\left(\dot{\xi}^d + \Gamma^d_{b\,c}\,\xi^b\dot{x}^c\right)\dot{x}^e = \frac{D^2\,\xi^a}{du^2} \quad ✓$$

The remaining terms can be relabeled so that they all end in $\xi^b \dot{x}^c\, \dot{x}^d$ :

$$-\Gamma^a_{de}\,\Gamma^d_{bc}\,\xi^b\dot{x}^c\dot{x}^e - \Gamma^a_{bc,d}\,\xi^b\dot{x}^c\dot{x}^d + \Gamma^a_{bc}\,\xi^b\,\Gamma^c_{de}\,\dot{x}^d\dot{x}^e + \Gamma^a_{bc,d}\,\dot{x}^b\dot{x}^c\xi^d$$

$$= (-\Gamma^a_{ed}\,\Gamma^e_{bc} - \Gamma^a_{bc,d} + \Gamma^a_{be}\,\Gamma^e_{cd} + \Gamma^a_{cd,b})\,\xi^b\dot{x}^c\dot{x}^d$$

$$\overset{(3.13)}{=} R^a{}_{cbd}\,\xi^b\dot{x}^c\dot{x}^d \qquad ✓$$

Thus, equation (3.4-5) can be simplified to read

$$\boxed{\frac{\mathrm{D}^2\,\xi^a}{\mathrm{d}u^2} + R^a{}_{cbd}\,\xi^b\,\dot{x}^c\,\dot{x}^d = 0}\ . \tag{3.35}$$

This is the **equation of geodesic deviation**.

For example, in a flat manifold, $R^a{}_{bcd} = 0$. In Euclidean space, $\frac{\mathrm{D}^2\,\xi^a}{\mathrm{d}u^2} = \frac{\mathrm{d}^2\,\xi^a}{\mathrm{d}u^2}$. Thus, the equation of geodesic deviation for flat Euclidean space is $\ddot{\xi}^a = 0$ and, hence, a parametric separation vector between two geodesics (i.e., lines) varies according to the formula $\xi^a = Au + B$. The two Euclidean geodesics deviate linearly with $u$, and with $s$ if $\gamma$ is non-null.

In a curved manifold, on the other hand, the separation between two geodesics accelerates with $u$. If acceleration is negative, the distance decreases until, possibly, the geodesics cross and then the distance increases, perhaps forever or perhaps until an inflection point is reached, upon which the distance decreases again.

A final point. How does one reconcile that gravity is a force in Newtonian physics with it not being a force, but curvature, in general relativity? The answer is that gravity is a “fictitious” force, just like the centrifugal, Euler, and Coriolis forces. Newton transformed away the latter forces by choosing a non-rotating reference frame centered at the particle. Had he chosen a reference frame that was freely falling as well, he could have transformed away the gravitational force.

## 3.5 Einstein's Field Equations

In 1914-15 Einstein made many attempts to find an equation that would relate matter and energy to spacetime curvature. Consider

$$g^{\mu\nu} = \kappa\, T^{\mu\nu} \tag{3.36}$$

where $\kappa$ is a constant. The metric tensor $g^{\mu\nu}$ is tied to curvature from the geodesic equation. The energy-momentum tensor $T^{\mu\nu}$ is defined in terms of the momentum density and pressure of an ideal fluid and, thus, involves all of the mass and energy information of the fluid. Both tensors are symmetric (i.e., they agree), and both tensors have zero divergence:

$$g^{\mu\nu}{}_{;\,\mu} \overset{(2.60)}{=} 0$$

and

$T^{\mu\nu}{}_{;\,\mu} = 0$ iff the relativity continuity equation (3.5) for conservation of energy and the relativity equation of motion ($f^{\mu} = \frac{\mathrm{d}p^{\mu}}{\mathrm{d}\tau}$) for a perfect fluid (3.6) are satisfied (by Theorem 3.1.1).

However, equation (3.36) does not reduce to **Poisson's equation**, the Newtonian field equation:

$$\nabla^2 V = 4\pi G\rho \tag{3.5-1}$$

where $V$ is the gravitational potential, $G$ is Newton's gravitational constant, and $\rho$ is the mass density of a point ($x^i$) under consideration.

In order to reduce to equation (3.5-1), the curvature tensor on LHS of equation (3.36) must have terms involving the 2nd derivative of $V$. To that end, in equation (2.83) of Section 2.7 we saw that

$$g_{00} = \eta_{00} + \frac{2\,V}{c^2},$$

meaning that $g_{00}$ has the same partial derivatives (up to a constant) as the gravitational potential $V$ in Euclidean space. In a more general curved manifold, the potential would get dispersed throughout $g_{\mu\nu}$. So, a tensor involving the 2nd derivative of $g^{\mu\nu}$ would be appropriate. One of Einstein's later attempts at the field equations was to use the Ricci tensor instead of $g^{\mu\nu}$:

$$R^{\mu\nu} = \kappa\, T^{\mu\nu}. \tag{3.37}$$

We observe that $R^{\mu\nu}$ involves 2nd derivatives of $g^{\mu\nu}$ because:

$$R^{\mu\nu} = g^{\mu\sigma} g^{\nu\rho} R_{\sigma\rho} \overset{(3.21)}{=} g^{\mu\sigma} g^{\nu\rho} R^{\chi}{}_{\sigma\rho\chi}$$

$$\overset{(3.13)}{=} g^{\mu\sigma} g^{\nu\rho} [\partial_\rho \Gamma^{\chi}_{\sigma\chi} - \partial_\chi \Gamma^{\chi}_{\sigma\rho} + \Gamma^{\zeta}_{\sigma\chi} \Gamma^{\chi}_{\zeta\rho} - \Gamma^{\zeta}_{\sigma\rho} \Gamma^{\chi}_{\zeta\chi}] ,$$

and

$$\Gamma^{\sigma}_{\mu\nu} \overset{(2.13)}{=} \frac{1}{2} g^{\sigma\rho} (\partial_\mu g_{\rho\nu} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})$$

so that $R^{\mu\nu}$ involves second derivatives of $g_{\mu\nu}$ and, consequently, of $g^{\mu\nu}$ since it is the inverse of $g_{\mu\nu}$.

$R^{\mu\nu}$ is also symmetric. However, $R^{\mu\nu}{}_{;\mu} \neq 0$.

Einstein later proposed

$$\boxed{G^{\mu\nu} \overset{(3.23)}{=} R^{\mu\nu} - \frac{1}{2} R g^{\mu\nu} = \kappa T^{\mu\nu}}, \tag{3.38}$$

which satisfies all of the requirements: it is symmetric, involves the metric tensor, reduces to Poisson's equation (as will be shown in the next section), and $G^{\mu\nu}{}_{;\mu} \overset{(\text{Th } 3.2.6)}{=} 0$. In reducing equations (3.38) to Poisson's equation, the constant $\kappa$ turns out to have the value - $8\pi \frac{G}{c^4}$. Observe that (3.38) consist of 10 equations:

$4\mu \times 4\nu = 16$ total cases. 6 duplicate cases $G_{\mu\nu} = G_{\nu\mu}$ ✓

Recapping what we have so far. If we are given a metric tensor $g^{\mu\nu}$, the corresponding Christoffel coefficients can be computed from it, and the Ricci tensor can be computed from the metic tensor and the Christoffel coefficients. So, $G$ is completely determined by a metric tensor or, equivalently, a line element.

Definition A **solution to Einstein's field equations** is a line element that satisfies equation (3.38).

Metric tensors are coordinate-dependent. If we change the coordinate system, the metric tensor might change. So, for example, the Schwarzschild solution, sometimes called the Schwarzschild geometry, is a coordinate system and a line element in terms of that coordinate system.

By Exercise (3.5.1), an alternate form for field equations (3.38) is

$$R^{\mu\nu} = \kappa\,(T^{\mu\nu} - \tfrac{1}{2}T\,g^{\mu\nu}) \tag{3.39}$$

where $T = T^{\mu}_{\mu}$. This makes it unclear whether there is a preference for the middle term to belong on the left (curvature) or on the right (mass-energy). If it is on the left, then both sides have zero divergence. If it is on the right side, then although neither side has zero divergence, they are still in agreement.

After the equations were discovered to show that the universe could not be static but must be either expanding or shrinking, Einstein added a **cosmological constant** term $\Lambda$, and later disavowed the term when the universe was found to be expanding. This is expressed in Chapter 6 as

$$R^{\mu\nu} - \tfrac{1}{2}R\,g^{\mu\nu} + \Lambda g^{\mu\nu} = \kappa T^{\mu\nu}\ . \tag{6.50}$$

Similar to $R = R^{\mu}_{\mu}$, we set $T = T^{\mu}_{\mu}$. Then, an alternate form for these field equations is

$$R^{\mu\nu} = \kappa\,(T^{\mu\nu} - \tfrac{1}{2}T\,g^{\mu\nu}) + \Lambda g^{\mu\nu}: \tag{3.5-2}$$

$$R^{\mu}_{\nu} - (\tfrac{1}{2}R - \Lambda)\,\delta^{\mu}_{\nu} = g_{\nu\sigma}\,[R^{\mu\sigma} - (\tfrac{1}{2}R - \Lambda)\,g^{\mu\sigma}] \overset{(3.38)}{=} g_{\nu\sigma}\,\kappa\,T^{\mu\sigma} = \kappa\,T^{\mu}_{\nu} \qquad \text{(a)}$$

$$\delta^{\mu}_{\mu} = 4 \qquad \text{(b)}$$

$$-R + 4\Lambda = R^{\mu}_{\mu} - (\tfrac{1}{2}R - \Lambda)\,(4) \overset{(b)}{=} R^{\mu}_{\mu} - (\tfrac{1}{2}R - \Lambda)\,\delta^{\mu}_{\mu} \overset{(a)}{=} \kappa\,T^{\mu}_{\mu} = \kappa T \qquad \text{(c)}$$

$$R \overset{(c)}{=} 4\Lambda - \kappa T \qquad \text{(d)}$$

$$R^{\mu\nu} \overset{(3.38)}{=} \kappa\,T^{\mu\nu} + (\tfrac{1}{2}R - \Lambda)\,g^{\mu\nu} \overset{(d)}{=} \kappa\,T^{\mu\nu} + (2\Lambda - \tfrac{1}{2}\kappa T - \Lambda)\,g^{\mu\nu}$$

$$= \kappa\,(T^{\mu\nu} - \tfrac{1}{2}T\,g^{\mu\nu}) + \Lambda\,g^{\mu\nu} \qquad \blacksquare$$

Field equations (6.50) and (3.5-2) can also be expressed in covariant terms:

$$R_{\mu\nu} - \tfrac{1}{2}R\,g_{\mu\nu} + \Lambda\,g_{\mu\nu} = \kappa T_{\mu\nu} \tag{3.5-3}$$

$$R_{\mu\nu} = \kappa\,(T_{\mu\nu} - \tfrac{1}{2}T\,g^{\mu\nu}) + \Lambda\,g_{\mu\nu} \tag{3.5-4}$$

$\Lambda g^{\mu\nu}$ allows equation (6.50) to be fine-tuned in order to match the rate of expansion (or contraction) found in cosmological measurements. If it is vacuum energy that drives the accelerated expansion of the universe, then $\Lambda$ may represent the vacuum energy. An unknown energy represented by $\Lambda$ is called **dark energy**. Whatever it represents, the new term is symmetric, has zero divergence since it is a constant times $g^{\mu\nu}$, and it disappears when reducing to Newtonian physics, so the equations still satisfy all of the requirements for general relativity field equations.

Current research is investigating whether dark energy may not be constant. If Lambda in equation (6.50) varies with time, $\Lambda(t)\, g^{\mu\nu}$ still has zero divergence because $\nabla \bullet (\Lambda(t) g^{\mu\nu}) = \Lambda(t)\, \nabla \bullet g^{\mu\nu}$, and so field equations (6.50) remain appropriate for modeling this phenomena.

$T^{\mu\nu}$ contains all forms of energy and momentum (other than dark energy, $\Lambda$). A region of spacetime in which $T^{\mu\nu} = 0$ is called a **vacuum** (or **empty**). It is devoid not only of matter but of radiative energy and momentum. From equation (3.39) the **vacuum field equations**, also called the **empty spacetime field equations**, are

$$R^{\mu\nu} = 0 \,. \tag{3.40}$$

The **Lambda vacuum field equations** (3.5-2) become

$$R^{\mu\nu} = \Lambda g^{\mu\nu} \tag{3.5-5}$$

which can also be expressed as

$$R_{\mu\nu} = \Lambda g_{\mu\nu} \,. \tag{3.5-6}$$

Focusing on the non-Lambda scenario, further support for the correctness of the field equations comes from comparing the equation of geodesic deviation with its Newtonian counterpart. With proper time $\tau$ as an affine parameter, equation (3.35) of geodesic deviation takes the form

$$\frac{\mathrm{D}^2 \xi^\mu}{\mathrm{d}\tau^2} + R^\mu{}_{\sigma\nu\rho}\, \xi^\nu \dot{x}^\sigma \dot{x}^\rho = 0 \tag{3.41}$$

where $\xi^\mu(\tau)$ is the small “vector” connecting corresponding points on neighboring geodesics.

For comparison with its Newtonian counterpart, let us write this as

$$\frac{\mathrm{D}^2 \xi^\mu}{\mathrm{d}\tau^2} = - K^\mu{}_\nu\, \xi^\nu \tag{3.42}$$

where

$$K^\mu{}_\nu = R^\mu{}_{\sigma\nu\rho}\, \dot{x}^\sigma \dot{x}^\rho \overset{(3.2\text{–}7)}{=} - R^\mu{}_{\sigma\rho\nu}\, \dot{x}^\sigma \dot{x}^\rho \,. \tag{3.43}$$

The corresponding situation in Newtonian theory is two particles moving under gravity on nearby paths $\tilde{x}^i(t)$ and $x^i(t)$. Equation (2.6-2) expresses acceleration in terms of potential:

$$a^i = -\partial_i V = -\delta^{ik}\,\partial_k V.$$

So, the equations of motion for the two particles are

$$\frac{d^2\tilde{x}^i}{dt^2} = -\delta^{ik}\,\tilde{\partial}_k V \quad \text{and} \quad \frac{d^2 x^i}{dt^2} = -\delta^{ik}\,\partial_k V \tag{3.5-7}$$

where

$\tilde{\partial}_k V$ is the gradient of the gravitational potential $V$ evaluated at $\tilde{x}^i(t)$.

Define "vectors" between the two paths:

$$\xi^i(t) \equiv \tilde{x}^i(t) - x^i(t). \tag{3.5-8}$$

The 1st order Taylor series approximation of $\tilde{\partial}_k V$ for small $\xi^i$ is (from equation 2.2-32)

$$\tilde{\partial}_k V \equiv f(\tilde{x}^k) \approx f(x^k) + \frac{\partial}{\partial x^j} f(\tilde{x}^k)\,(\tilde{x}^k - x^k) = \partial_k V + (\partial_j\partial_k V)\,\xi^j\,. \tag{3.5-9}$$

So,

$$\frac{d^2\xi^i}{dt^2} = -K^i_j\,\xi^j \tag{3.44}$$

where

$$K^i_j \equiv \delta^{ik}\,\partial_j\partial_k V: \tag{3.45}$$

$$\frac{d^2\xi^i}{dt^2} \overset{(3.5-8)}{=} \frac{d^2\tilde{x}^i}{dt^2} - \frac{d^2 x^i}{dt^2} \overset{(3.5-7)}{=} -\delta^{ik}(\tilde{\partial}_k V - \partial_k V) \overset{(3.5-9)}{=} -\delta^{ik}\,\partial_j\partial_k V\,\xi^j = -K^i_j\,\xi^j \ \checkmark$$

Equation (3.44) is the Newtonian counterpart of geodesic deviation equation (3.42), which shows that

$$K^\mu{}_\nu = -R^\mu{}_{\sigma\rho\nu}\,\dot{x}^\sigma\dot{x}^\rho \qquad \longleftrightarrow \qquad K^i_j = \delta^{ik}\,\partial_j\partial_k V\,.$$

The **empty space field equation of Newtonian gravitation** is $\nabla^2 V \overset{(2.6\text{–}4)}{=} 0$. This is equivalent to $K^i_i = 0$:

$$\nabla^2 V \equiv (\nabla \bullet \nabla) V = \frac{\partial^2 V}{(\partial x^1)^2} + \frac{\partial^2 V}{(\partial x^2)^2} + \frac{\partial^2 V}{(\partial x^3)^2} = \sum \partial_i^2 V$$

$$K^i_i = K^i_\ell \, \delta^\ell_i = \delta^\ell_i \, \delta^{ik} \, \partial_\ell \partial_k V = \delta^{ik} \, \partial_i \partial_k V = \sum \partial_i \partial_i V = \sum \partial_i^2 V$$

$$\Rightarrow \quad \nabla^2 V = K^i_i \qquad \checkmark$$

Thus, the general relativity vacuum field equations (3.40) are the counterpart to the Newtonian empty space field equation:

Newtonian empty space field equation $\Leftrightarrow \ \nabla^2 V = 0 \ \Leftrightarrow \ K^i_i = 0$

$\longleftrightarrow \ K^\mu{}_\mu = 0 \ \Leftrightarrow \ R_{\sigma\rho} = R^\mu{}_{\sigma\rho\mu} = 0 \ \Leftrightarrow \ R^{\mu\nu} = g^{\mu\sigma} g^{\nu\rho} R_{\sigma\rho} = 0$

$\Leftrightarrow$ general relativity vacuum field equations $\checkmark$

This lends support for for the vacuum field equations. Support for the vacuum field equations (6.50) and (3.5-2), of which (3.38) and (3.39) are special cases, is given in the next section.

## 3.6 Einstein’s Equation compared with Poisson’s equation

The field equation of Newtonian gravitation is Poisson’s equation:

$$\nabla^2 V = 4\pi G\rho \,. \qquad (2.6.4)$$

This represents one equation in the one unknown, $V$. By equation (2.6.1),

$$V = - \int_{-\infty}^{x} g \, \mathrm{d}x.$$

So, solving for $V$ in equation (2.6.4) is equivalent to solving for $g$.

The covariant version of Einstein’s gravitational field equations is

$$R_{\mu\nu} = \kappa \left(T_{\mu\nu} - \tfrac{1}{2} T \, g_{\mu\nu}\right) :$$

$$R_{\mu\nu} = g_{\mu\sigma} \, g_{\nu\rho} \, R^{\sigma\rho} \overset{(3.39)}{=} g_{\mu\sigma} \, g_{\nu\rho} \, \kappa \left(T^{\sigma\rho} - \tfrac{1}{2} T \, g^{\sigma\rho}\right) = \kappa \left(T_{\mu\nu} - \tfrac{1}{2} T \, g_{\mu\nu}\right) \checkmark$$

Because $g_{\mu\nu} = g_{\nu\mu}$, we can restrict our attention to $\mu \leq \nu$, reducing the 16 gravitational field equations in 16 unknowns down to 10 equations in 10 unknowns, as compared to the Newtonian equation (2.6.4) that is only 1 equation in 1 unknown.

To show that Einstein’s field equations reduce to Poisson’s equation, we will show that $R_{00}$ reduces to Poisson’s equation under approximately Euclidean conditions.

$$R_{00} = \kappa \left(T_{00} - \tfrac{1}{2} T \, g_{00}\right) . \qquad (3.46)$$

Not mentioned in the book, we must, in addition, show that the other 9 simultaneous equations are redundant. To that end, in the next section, Section 3.7, when we develop the Schwarzschild solution to Einstein’s field equations, we will show that solving for the non-zero connection coefficients reduces the ten field equations down to four equations (3-54 to 3-57) in two unknowns, $A$ and $B$. We then show that in asymptotically flat space that $B \propto 1/A$, reducing this to four equations in one unknown, which are then shown to be consistent under the approximately Newtonian conditions assumed by Schwarzschild. Thus, the ten simultaneous field equations reduce to the single
equation (3.46) ⇔ equation (3-54) as Newtonian conditions are approached.

To show that $R_{00}$ reduces to equation (3.46), as in Section 2.7, we assume we have an almost Cartesian coordinate system in which:

- $g_{\mu\nu} \overset{(2.74)}{=} \eta_{\mu\nu} + h_{\mu\nu}$ where $h_{\mu\nu}$ are small (i.e., $h_{\mu\nu} \approx 0$ to 1st order) (3.6-1)

- Products

$$h_{\mu\nu}\, h_{\sigma\rho} \approx h^{\mu\nu}\, h^{\sigma\rho} \approx h^{\mu\nu} h_{\sigma\rho} \approx 0 \text{ to 2nd order} \tag{3.6-2}$$

- Weak field products

$$\partial_\alpha h_{\mu\nu}\, \partial_\beta h_{\sigma\rho} \approx 0 \;\text{ to 2nd order} \tag{3.6-3}$$

- The gravitational field $h_{\mu\nu}$ is quasi-static:

$$|\partial_0 h_{\mu\nu}| \overset{(2.7-7)}{\ll} |\partial_i h_{\sigma\rho}| \;\text{ and }\; |\partial_0 \Gamma^\mu_{\nu\alpha}| \ll |\partial_\mathrm{i} \Gamma^\sigma_{\rho\beta}| \text{ for all } i, \alpha, \beta, \mu, \nu, \sigma, \rho\,. \tag{3.6-4}$$

The quasi-static condition (3.6-4) is a much stronger version than previously given as equation (2.7-7). In thermodynamics, quasi-static means that changes happen slowly so that equilibrium is maintained. In this section we need not only that 1st derivatives change slowly (1st order) but also that 2nd derivatives change slowly (2nd order). The second part of equations (3.6-4) includes 2nd derivative terms like $\partial_0 \partial_\gamma h_{\nu\alpha} \ll \partial_i \partial_\delta h_{\rho\beta}$. In fact, requiring that $\Gamma^\mu_{\nu\alpha}$ change slowly is an even more stringent requirement than just that 2nd derivatives change slowly.

We also assume

- Particles (in this coordinate system) have speeds $v \ll c$ so that

$$\gamma \approx 1 \;\text{ and }\; \left| \frac{\mathrm{dx}^i}{\mathrm{dt}} \right| \overset{(2.7-6)}{\ll} c\,. \tag{3.6-5}$$

From Appendix A.0, Example 1,

$$u_0 \approx c \;\text{ and }\; |u_i| \ll c: \tag{3.6-6}$$

$$u^\mu = \gamma\,(c, \boldsymbol{v}) \;\Rightarrow\; u_0 = u^0 \overset{(3.6-5)}{\approx} c$$

$$|u_i| = |-u^i| = |\gamma v^i| \overset{(3.6-5)}{\approx} |v^i| \overset{(A.23)}{=} \left|\frac{\mathrm{dx}^i}{\mathrm{dt}}\right| \overset{(3.6-5)}{\ll} c. \checkmark$$

Finally, we also assume

- $$\frac{P}{c^2} \ll \rho\,, \text{ or } P \ll \rho c^2 \tag{3.6-7}$$

  as is true for most classical distributions (Sun, water, high pressure gas).

Under these conditions, the covariant version of the energy-momentum stress tensor becomes $T_{\mu\nu} \approx \rho\, u_\mu u_\nu - P g_{\mu\nu}$ :

$$T_{\mu\nu} \overset{(3.1-19)}{=} (\rho + \frac{P}{c^2})\, u_\mu u_\nu - P g_{\mu\nu} \overset{(3.6-7)}{\approx} \rho\, u_\mu u_\nu - P g_{\mu\nu}$$

Thus, $T_{00} \approx \rho c^2$ and $T_{i\,i} \approx \rho\,(u_i)^2 + P$ : (3.6-8)

$$T_{00} \approx \rho\,(u_0)^2 - P\,\eta_{00} = \rho\,(u_0)^2 - P \overset{(3.6-6)}{\approx} \rho c^2 - P \overset{(3.6-7)}{\approx} \rho c^2$$

$$T_{i\,i} \approx \rho\,(u_i)^2 - P\,\eta_{i\,i} = \rho\,(u_i)^2 + P$$

Claim $T \approx \rho c^2$ : (3.6-9)

$$T = T^\mu_\mu = g^{\mu\nu} T_{\mu\nu} = g^{00} T_{00} + g^{i\,i} T_{i\,i} \overset{(3.6-8)}{\approx} (1)\,\rho c^2 + (-1)\,[\rho\,(u_i)^2 + P]$$

$$= \rho\, c^2 - \rho\, u_i{}^2 - P \overset{(3.6-7)}{\approx} \rho c^2$$

Claim $R_{00} \approx \frac{1}{2}\kappa\rho c^2$ : (3.47)

$$R_{00} \overset{(3.46)}{=} \kappa\,(T_{00} - \tfrac{1}{2} T\, g_{00}) \overset{(3.6-8,\, 3.6-9)}{\approx} \kappa\rho\,(u_0{}^2 - \tfrac{1}{2}\, c^2\, g_{00})$$

$$\overset{(3.6-6)}{\approx} \kappa\rho\,[c^2 - \tfrac{1}{2} c^2 (1)] = \tfrac{1}{2}\,\kappa\rho c^2 \qquad \checkmark$$

Also,

$$R_{00} \overset{(3.21)}{=} R^\mu_{00\,\mu} \overset{(3.13)}{=} \partial_0 \Gamma^\mu_{0\,\mu} - \partial_\mu \Gamma^\mu_{00} + \Gamma^\nu_{0\,\mu}\,\Gamma^\mu_{\nu 0} - \Gamma^\nu_{00}\,\Gamma^\mu_{\nu\mu} \qquad (3.48)$$

Claim $\Gamma^\mu_{\nu\alpha}\,\Gamma^\sigma_{\rho\beta} \approx 0$ : (3.6-10)

From Section 2.7 we have that

$$\Gamma^\mu_{\nu\alpha}\,\Gamma^\sigma_{\rho\beta} \overset{(2.78)}{\approx} \frac{1}{4}\,\eta^{\mu\gamma}\eta^{\nu\delta}\,(\partial_\nu h_{\gamma\alpha} + \partial_\alpha h_{\nu\gamma} - \partial_\gamma h_{\nu\alpha})\,(\partial_\rho h_{\delta\beta} + \partial_\beta h_{\rho\delta} - \partial_\delta h_{\rho\beta}) \overset{(3.6-3)}{\approx} 0 \quad \checkmark$$

By (3.6-4), $|\partial_0 \Gamma^\mu_{\nu\alpha}| \ll |\partial_{\mathrm{i}} \Gamma^\sigma_{\rho\beta}|$ for all $i$, $\alpha$, $\beta$, $\mu$, $\nu$, $\sigma$, $\rho$ . Thus equation (3.48) simplifies to

$$R_{00} = -\,\partial_{\mathrm{i}} \Gamma^{\mathrm{i}}_{00}\,. \qquad (3.6-11)$$

Claim $R_{00} \approx -\frac{1}{c^2}\,\nabla^2 V$ : (3.6-12)

$$R_{00} \overset{(3.6-11)}{=} -\,\partial_{\mathrm{i}}\Gamma^{\mathrm{i}}_{00} = -\,(\partial_1 \Gamma^1_{00} + \partial_2 \Gamma^2_{00} + \partial_3 \Gamma^3_{00})$$

$$\overset{(2.7-8)}{=} -\,\frac{1}{2}(\partial_1\partial_1 h_{00} + \partial_2\partial_2 h_{00} + \partial_3\partial_3 h_{00})$$

$$= -\,\frac{1}{2}\,\nabla^2 h_{00} \overset{(2.83)}{=} -\,\frac{1}{2}\,\nabla^2\,(\frac{2\,V}{c^2})$$

$$= -\frac{1}{c^2}\,\nabla^2 V \quad ✓$$

$$\therefore\ \nabla^2 V \approx -\tfrac{1}{2}\kappa\rho c^4 : \tag{3.49}$$

$$\nabla^2 V \overset{(3.6-12)}{\approx} -c^2 R_{00} \overset{(3.47)}{\approx} -\tfrac{1}{2}\kappa\rho c^4 ✓$$

Equation (3.49) is precisely Poisson's equation (2.6.4) if

$$\boxed{\kappa = -\frac{8\,\pi\,G}{c^4}} : \tag{3.6-13}$$

$$\nabla^2 V \approx 4\pi G\rho \quad ✓ \tag{3.50}$$

This shows that $c^2\, R_{00} \approx 4\pi G\rho$; that is, the 1st of the ten $R_{\mu\nu} = \kappa\,(T_{\mu\nu} - \frac{1}{2} T\, g_{\mu\nu})$ field equations is precisely Poisson's equation.

## 3.7 The Schwarzschild Solution

Schwarzschild examined the special case of a spherically symmetric gravitational field in the empty spacetime surrounding a massive spherical object like a star. His assumptions are:

(a) The metric tensor field is static
(b) The metric tensor field is spherically symmetric
(c) Spacetime is empty
(d) Spacetime is asymptotically flat
(e) Spacetime can be coordinatized by $(t, r, \theta, \phi)$ where **coordinate *t* is timelike**, meaning the tangent vector to the coordinate curve is timelike, $\theta$ and $\phi$ are the usual spherical angles, and $r$ is a radial coordinate.

He postulated a line element

$$c^2\, d\tau^2 = A(r)\, dt^2 - B(r)\, dr^2 - r^2\, d\theta^2 - r^2 \sin^2\theta\, d\phi^2\ . \tag{3.51}$$

that can also be expressed as

$$\begin{aligned} c^2\, d\tau^2 &= \tfrac{1}{c^2} A(r)\, d(c\,t)^2 - B(r)\, dr^2 - r^2\, d\theta^2 - r^2 \sin^2\theta\, d\phi^2 \\ &= g_{00}\, d(c\,t)^2 + g_{11}\, dr^2 + g_{22}\, d\theta^2 + g_{33}\, d\phi^2. \end{aligned} \tag{3.7-1}$$

$$\Rightarrow \quad (g_{\mu\nu}) = \begin{pmatrix} \frac{1}{c^2} A(r) & & & \\ & -B(r) & & \\ & & -r^2 & \\ & & & -r^2 \sin^2\theta \end{pmatrix} \quad \forall\ r, \tag{3.7-2}$$

and

$$(g^{\mu\nu}) = \begin{pmatrix} c^2 / A(r) & & & \\ & -1/B(r) & & \\ & & -1/r^2 & \\ & & & -1/r^2 \sin^2\theta \end{pmatrix}, \ r > 0. \tag{3.7-3}$$

Assumption (a) is satisfied because all $g_{\mu\nu}$ are independent of $t$. ✓

If we freeze $r$ and $t$, we get a surface $\Sigma$ whose line element is

$$ds^2 = r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2 \tag{3.52}$$

By Exercise 1.6.2 (a), this is the line element of a sphere of radius *r*, parameterized by $\theta$ and $\phi$. This expresses assumption (b). ✓

Assumption (c) means that $A(r)$ and $B(r)$ are to be solved using the vacuum field equations (3.5-6): $R_{\mu\nu} = 0$.

Assumption (d) gives the boundary conditions on *A* and *B*, namely

$$A(r) \to c^2 \quad \text{and} \quad B(r) \to 1 \quad \text{as} \quad r \to \infty. \tag{3.53}$$

Here is an outline of the steps we take to solve the vacuum field equations to determine $A(r)$ and $B(r)$.

- We first note that the only non-zero metric tensor components are the four components on the diagonal of the matrix
- This results in only thirteen non-zero connection coefficients (left table, below).
- In turn, this leads to only four non-zero Ricci tensors $R_{\mu\nu}$, formulas (3.54) – (3.57), below.
- Thus, the original ten "$R_{\mu\nu} = 0$" vacuum field equations are reduced to only four equations, and in only two unknowns, $A(r)$ and $B(r)$.
- The first two *differential* equations can be combined into a single differential equation of $A(r)$ and $B(r)$ that results in $B(r) = \dfrac{c^2}{A(\mathrm{r})}$.
- We observe that the fourth equation is is a multiple of the 3rd equation and, thus, redundant.
- When we replace $B(r)$ by $\dfrac{c^2}{A(\mathrm{r})}$ in either of the first two equations, we get a new *linear* equation in $A(r)$ that does not yield a meaningful solution because the derivatives $A'[r]$ and $A''[r]$ factor out, meaning we lose necessary constants of integration.
- However, the 3rd equation yields a formula for $A(r)$ having a new unknown, *k*, a constant of integration.
  - ✓ Plugging this formula into the first two differential equations confirms that they are consistent with this result.
- We next solve for *k* by performing a change of coordinates between Euclidean and Schwarzschild spacetime coordinates, and then using Euclidean properties to generate one equation in the one unknown, *k*.
  - ✓ This solves for $A(r)$, hence $B(r)$, and hence $g_{\mu\nu}$.

These steps were referenced in Section 3.6 where it was shown that Einstein's ten field equations reduce to Poisson's one equation in Euclidean space.

We now execute the outline above. For easier readability while solving the field equations, we shorten $A(r)$ and $B(r)$ to $A$ and $B$, respectively, and we use primes to denote derivatives with respect to $r$:

We start by computing the non-zero connection coefficients $\Gamma^{\mu}_{\nu\gamma}$ (table, below). Next, using equations (3.21) and (3.13) we get

$$R_{\mu\nu} \overset{(3.21)}{=} R^{\sigma}_{\ \mu\nu\sigma} \overset{(3.13)}{\equiv} \partial_{\nu}\Gamma^{\sigma}_{\mu\sigma} - \partial_{\sigma}\Gamma^{\sigma}_{\mu\nu} + \Gamma^{\rho}_{\mu\sigma}\Gamma^{\sigma}_{\rho\nu} - \Gamma^{\rho}_{\mu\nu}\Gamma^{\sigma}_{\rho\sigma}. \quad (3.7\text{-}4)$$

which is used to compute the non-zero Ricci tensors $R_{\mu\nu}$ (for $r > 0$), where prime ($'$) represents partial derivative with respect to $r$:

| $\mu$ | $\nu$ | $\gamma$ | $\Gamma^{\mu}_{\nu\gamma}$ |
|---|---|---|---|
| 0 | 0 | 1 | $\frac{A'}{2A}$ |
| 0 | 1 | 0 | $\frac{A'}{2A}$ |
| 1 | 0 | 0 | $\frac{A'}{2Bc^2}$ |
| 1 | 1 | 1 | $\frac{B'}{2B}$ |
| 1 | 2 | 2 | $-\frac{r}{B}$ |
| 1 | 3 | 3 | $-\frac{r\,\mathrm{Sin}[\theta]^2}{B}$ |
| 2 | 1 | 2 | $\frac{1}{r}$ |
| 2 | 2 | 1 | $\frac{1}{r}$ |
| 2 | 3 | 3 | $-\mathrm{Cos}[\theta]\,\mathrm{Sin}[\theta]$ |
| 3 | 1 | 3 | $\frac{1}{r}$ |
| 3 | 2 | 3 | $\mathrm{Cot}[\theta]$ |
| 3 | 3 | 1 | $\frac{1}{r}$ |
| 3 | 3 | 2 | $\mathrm{Cot}[\theta]$ |

| $\mu$ | $\nu$ | $R_{\mu\nu}$ | |
|---|---|---|---|
| 0 | 0 | $-\frac{A'}{Bc^2r} + \frac{(A')^2}{4ABc^2} + \frac{A'B'}{4B^2c^2} - \frac{A''}{2Bc^2}$ | (3.54) |
| 1 | 1 | $-\frac{(A')^2}{4A^2} - \frac{B'}{Br} - \frac{A'B'}{4AB} + \frac{A''}{2A}$ | (3.55) |
| 2 | 2 | $-1 + \frac{1}{B} + \frac{rA'}{2AB} - \frac{rB'}{2B^2}$ | (3.56) |
| 3 | 3 | $-\mathrm{Sin}[\theta]^2 + \frac{\mathrm{Sin}[\theta]^2}{B} + \frac{r\,\mathrm{Sin}[\theta]^2A'}{2AB} - \frac{r\,\mathrm{Sin}[\theta]^2B'}{2B^2}$ | (3.57) |

(Generation of $\Gamma^{\mu}_{\nu\gamma}$ constitutes Problem 2.7. Generation of (3.54 – 3.57) constitutes Exercise 3.7.1.)

Because of Assumption (c), we wish to solve the four vacuum field equations $R_{\mu\nu} = 0$ in the two unknowns $A$ and $B$. We observe that $R_{33} \overset{(3.56,\,3.57)}{=} \sin^2\theta\, R_{22}$. That is, $R_{33} = 0$ reduces to $R_{22} = 0$, which makes the fourth equation redundant.

$R_{00} = 0$ and $R_{11} = 0$ constitute two differential equations in the two unknowns $A$ and $B$ that we combine and solve as follows:

$$0 = \frac{B}{A} R_{00} + R_{11}$$

$$\overset{(3.54,\,3.55)}{=} \left[-\frac{A'}{A\,r} + \frac{(A')^2}{4\,A^2} + \frac{A'\,B'}{4\,A\,B} - \frac{A''}{2\,A}\right] + \left[-\frac{B'}{B\,r} - \frac{(A')^2}{4\,A^2} - \frac{A'\,B'}{4\,A\,B} + \frac{A''}{2\,A}\right]$$

$$= -\frac{1}{r}\left(\frac{A'}{A} + \frac{B'}{B}\right) = -\frac{1}{r}\,\frac{A'\,B + A\,B'}{A\,B} = -\frac{(A\,B)'}{r\,A\,B}$$

$\Rightarrow (A\,B)' = 0 \;\;\Rightarrow\;\; AB = K$, where $K$ is a constant

$\Rightarrow \;\; c^2\,(1) \overset{(3.53)}{=} \mathrm{Lim}_{r\to\infty}\, AB = \mathrm{Lim}_{r\to\infty}\, K$

$\Rightarrow \;\; K = c^2$

$$\Rightarrow \;\; B = \frac{c^2}{A}. \tag{3.7-5}$$

Using this, we see that $R_{22} = 0$ now becomes the one equation, (3.7-6), in the single unknown, $A$.

$$-1 + \frac{A}{c^2} + \frac{r\,A'}{c^2} = 0: \tag{3.7-6}$$

$$B' \overset{(3.7-5)}{=} -c^2\,\frac{A'}{A^2} \tag{3.7-7}$$

$$R_{22} \overset{(3.56)}{=} -1 + \frac{1}{B} + \frac{r\,A'}{2\,A\,B} - \frac{r\,B'}{2\,B^2} \overset{(3.7-5,\,3.7-7)}{=} -1 + \frac{A}{c^2} + \frac{r\,A'}{2\,c^2} - \frac{r}{2}\left(-\frac{c^2\,A'}{A^2}\right)\frac{A^2}{c^4}$$

$$= -1 + \frac{A}{c^2} + \frac{r\,A'}{2\,c^2} + \frac{r\,A'}{2\,c^2} = -1 + \frac{A}{c^2} + \frac{r\,A'}{c^2} \qquad ✓$$

$\Rightarrow \quad (r\,A)' = r\,A' + A \overset{(3.7-6)}{=} c^2$

$\Rightarrow \quad r\,A = \int c^2\, dr = c^2\,(r + k)$

$$\Rightarrow \quad A = c^2\left(1 + \frac{k}{r}\right) \tag{3.7-8}$$

$$\Rightarrow \quad B \overset{(3.7-5)}{=} \left(1 + \frac{k}{r}\right)^{-1} \tag{3.7-9}$$

Plugging $A$ and $B$ into $R_{00} = 0$ and $R_{11} = 0$ yield identities, so they are consistent but can't be used to solve for $k$. This is fortunate because $A$ needs to reflects mass or gravity, and so should include Newton's gravitational constant $G$. We find the value of $A$ by imposing condition (d), that spacetime be asymptotically flat, along with the aforementioned change of coordinates.

Substituting the expressions (3.7-8) and (3.7-9) for $A(r)$ and $B(r)$ into line element equation (3.51) yields

$$c^2\,\tau^2 = c^2\left(1+\frac{k}{r}\right)dt^2 - \left(1+\frac{k}{r}\right)^{-1}dr^2 - r^2 d\theta^2 - r^2 sin^2\theta\, d\phi^2 \qquad (3.7\text{-}10)$$

As $r \to \infty$, the line element approaches

$$c^2\,\tau^2 = c^2\,dt^2 - dr^2 - r^2\,d\theta^2 - r^2\,sin^2\theta\,d\phi^2, \qquad (3.7\text{-}11)$$

which is the line element for spherical coordinates in flat spacetime. Next, we perform change-of-coordinates, choosing lower-case indices for Schwarzschild coordinates:

$$x^0 = c\,t,\quad x^1 = r,\quad x^2 = \theta,\quad x^3 = \phi, \qquad (3.7\text{-}12)$$

where, in flat space,

$$x = r\sin\theta\cos\phi,\;\; y = r\sin\theta\sin\phi,\;\; z = r\cos\theta. \qquad (3.7\text{-}13)$$

For flat spacetime Cartesian coordinates, we choose upper case:

$$X^0 = c\,t,\quad X^1 = x,\quad X^2 = y,\quad X^3 = z. \qquad (3.7\text{-}14)$$

Recall

$$r = \sqrt{x^2+y^2+z^2},\quad \theta = \text{arc cos}\,\frac{z}{r},\quad \phi = \text{arc tan}\,\frac{y}{x}. \qquad (3.7\text{-}15)$$

From equation (3.7-10) we capture the four non-zero Schwarzschild-coordinates covariant metric tensors $g_{\mu\nu}$ and then express them in Cartesian spacetime coordinates:

$$\begin{aligned} g_{00} &= 1+\frac{k}{r} = 1+\frac{k}{\sqrt{x^2+y^2+z^2}} \\ g_{11} &= -\left(1+\frac{k}{r}\right)^{-1} = -\left[1+\frac{k}{\sqrt{x^2+y^2+z^2}}\right]^{-1} \\ g_{22} &= -r^2 = -(x^2+y^2+z^2) \\ g_{33} &= -r^2\,sin^2\theta = -r^2\,(1-\cos^2\theta) = -r^2+z^2 = -(x^2+y^2) \end{aligned} \qquad (3.7\text{-}16)$$

To find the Cartesian metric tensors $G_{MN}$ , we must perform a change of coordinates operation. Both Schwarzschild and Euclidean coordinate systems represent inertial frames because they each satisfy the inertial-frame definition that their respective coordinates are orthogonal. For example, $\frac{\partial r}{\partial \theta} = 0$ and $\frac{\partial x}{\partial t} = 0$ . Thus, Equation (A.8), $G_{MN} = \Lambda^{\sigma}_{M}\,\Lambda^{\rho}_{N}\,g_{\sigma\rho}$ , the Lorentz transformation between inertial frames, is the formula to use for changing Schwarzschild coordinates to Cartesian coordinates.

To use this equation, we must find an expression for $\Lambda^{\sigma}_{M}$. Recall from equation (A1.2) that the homogeneous Lorentz transformation $\tilde{\Lambda}^{\sigma}_{M}$ has the formula $\tilde{\Lambda}^{\sigma}_{M} = \frac{\partial x^{\sigma}}{dX^{M}}$, and that the (general) Lorentz transformation $\Lambda^{\sigma}_{M} = \tilde{\Lambda}^{\sigma}_{M}$ iff the offset $a^{\mu} = 0$. Since Schwarzschild coordinates have the same origin as Cartesian spacetime coordinates, the offset is zero. So,

$$\Lambda^{\sigma}_{M} = \frac{\partial x^{\sigma}}{\partial X^{M}}, \tag{3.7-17}$$

and we have

$$G_{MN} = \Lambda^{\sigma}_{M}\,\Lambda^{\rho}_{N}\,g_{\sigma\rho} = \frac{\partial x^{\sigma}}{\partial X^{M}}\frac{\partial x^{\rho}}{\partial X^{N}}\,g_{\sigma\rho}\,. \tag{3.7-18}$$

Since $g_{\sigma\rho} = 0$ unless $\sigma = \rho$, this simplifies to $G_{MN} = \Lambda^{\sigma}_{M}\,\Lambda^{\sigma}_{N}\,g_{\sigma\sigma}$ , and expanding this out yields

$$G_{MN} = \Lambda^{0}_{M}\,\Lambda^{0}_{N}\,g_{00} + \Lambda^{1}_{M}\,\Lambda^{1}_{N}g_{11} + \Lambda^{2}_{M}\,\Lambda^{2}_{N}g_{22} + \Lambda^{3}_{M}\,\Lambda^{3}_{N}g_{33} \tag{3.7-19}$$

We break this into three pieces:

$$G_{00} = \frac{\partial(c\,t)}{\partial(c\,t)}\frac{\partial(c\,t)}{\partial(c\,t)}\,g_{00} + 0 + 0 + 0 \overset{(3.7\text{–}16)}{=} 1 + \frac{k}{r} \tag{3.7-20}$$

$$\begin{aligned} G_{0I} &= \frac{\partial(c\,t)}{\partial(c\,t)}\frac{\partial(\text{ct})}{\partial x^{I}}\,g_{00} + \frac{\partial x^{1}}{\partial(c\,t)}\frac{\partial x^{1}}{\partial x^{I}}\,g_{11} + \frac{\partial x^{2}}{\partial(c\,t)}\frac{\partial x^{2}}{\partial x^{I}}\,g_{22} + \frac{\partial x^{3}}{\partial(c\,t)}\frac{\partial x^{3}}{\partial x^{I}}\,g_{33} \\ &= 0 \quad \left(\text{because } \frac{\partial(c\,t)}{\partial x^{I}} = 0 \text{ and } \frac{\partial x^{i}}{\partial(c\,t)} = 0\right) \end{aligned} \tag{3.7-21}$$

$$G_{IJ} = \frac{\partial x^{1}}{\partial x^{I}}\frac{\partial x^{1}}{\partial x^{J}}\,g_{11} + \frac{\partial x^{2}}{\partial x^{I}}\frac{\partial x^{2}}{\partial x^{J}}\,g_{22} + \frac{\partial x^{3}}{\partial x^{I}}\frac{\partial x^{3}}{\partial x^{J}}\,g_{33} \tag{3.7-22}$$

We list the values of $\Lambda^{\sigma}_{M}$, based on equations (3.7-17 & 3.7-12 to -14), in Table 3.7-1, below. We list the values of $G_{M\,N}$ for $r > 0$, based on formulas (3.7-20, -21, & -22), in Table 3.7-2.

TABLE 3.7-1

| $\sigma$ | $M$ | $\Lambda^{\sigma}_{M} = \frac{\partial x^{\sigma}}{\partial X^{M}}$ |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 3 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | $\frac{x}{r}$ |
| 1 | 2 | $\frac{y}{r}$ |
| 1 | 3 | $\frac{z}{r}$ |
| 2 | 0 | 0 |
| 2 | 1 | $\frac{x\,z}{r^2\sqrt{x^2+y^2}}$ |
| 2 | 2 | $\frac{y\,z}{r^2\sqrt{x^2+y^2}}$ |
| 2 | 3 | $-\frac{\sqrt{x^2+y^2}}{r^2}$ |
| 3 | 0 | 0 |
| 3 | 1 | $-\frac{y}{x^2+y^2}$ |
| 3 | 2 | $\frac{x}{x^2+y^2}$ |
| 3 | 3 | 0 |

TABLE 3.7-2

| $M$ | N | $G_{M\,N} = \Lambda^{\sigma}_{M}\,\Lambda^{\sigma}_{N}\,g_{\sigma\,\sigma}$ |
|---|---|---|
| 0 | 0 | $1+\frac{k}{r}$ |
| 0 | 2 | 0 |
| 0 | 2 | 0 |
| 0 | 3 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | $-1+\frac{k\,x^2}{r^2\,(r+k)}$ |
| 1 | 2 | $-\frac{k\,x\,y}{r^2\,(r+k)}$ |
| 1 | 3 | $-\frac{k\,x\,z}{r^2\,(r+k)}$ |
| 2 | 0 | 0 |
| 2 | 1 | $-\frac{k\,x\,y}{r^2\,(r+k)}$ |
| 2 | 2 | $-1+\frac{k\,y^2}{r^2\,(r+k)}$ |
| 2 | 3 | $-\frac{k\,y\,z}{r^2\,(r+k)}$ |
| 3 | 0 | 0 |
| 3 | 1 | $-\frac{k\,x\,z}{r^2\,(r+k)}$ |
| 3 | 2 | $-\frac{k\,y\,z}{r^2\,(r+k)}$ |
| 3 | 3 | $-1+\frac{k\,z^2}{r^2\,(r+k)}$ |

(3.7-23)

These tables constitute Exercise 3.7.2.

As $r\to\infty$, the Schwarzschild geometry approaches flat spacetime, and the Cartesian metric tensor $G$ approaches $\eta + h$, where $h = 0$ to the 2nd order. That is,

$G_{MN} = \eta_{MN} + h_{MN}$:

$$G_{MN} = \begin{pmatrix} 1+\frac{k}{r} & 0 & 0 & 0 \\ 0 & -1+\frac{k\,x^2}{r^2\,(r+k)} & -\frac{k\,x\,y}{r^2\,(r+k)} & -\frac{k\,x\,z}{r^2\,(r+k)} \\ 0 & -\frac{k\,x\,y}{r^2\,(r+k)} & -1+\frac{k\,y^2}{r^2\,(r+k)} & -\frac{k\,y\,z}{r^2\,(r+k)} \\ 0 & -\frac{k\,x\,z}{r^2\,(r+k)} & -\frac{k\,y\,z}{r^2\,(r+k)} & -1+\frac{k\,z^2}{r^2\,(r+k)} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} + \begin{pmatrix} \frac{k}{r} & 0 & 0 & 0 \\ 0 & \frac{kx^2}{r^2\,(r+k)} & -\frac{kx\,y}{r^2\,(r+k)} & -\frac{kx\,z}{r^2\,(r+k)} \\ 0 & -\frac{kx\,y}{r^2\,(r+k)} & \frac{ky^2}{r^2\,(r+k)} & -\frac{ky\,z}{r^2\,(r+k)} \\ 0 & -\frac{kx\,z}{r^2\,(r+k)} & -\frac{ky\,z}{r^2\,(r+k)} & \frac{kz^2}{r^2\,(r+k)} \end{pmatrix} \tag{3.7-24}$$

This allows us to solve for the constant $k$ (for $r > 0$):

$$\frac{k}{r} = -\frac{2\,M\,G}{c^2\,r}: \tag{3.7-25}$$

$$\frac{k}{r} \overset{(3.7\text{-}24)}{=} h_{00} \overset{(2.83)}{=} \frac{2\,V}{c^2} \overset{(2.8\text{-}2)}{=} -\frac{2}{c^2}\,\frac{GM}{r} \quad ✓$$

Replacing $\frac{k}{r}$ in the line element (3.7-10) generates the Schwarzschild line element and corresponding metric tensor:

$$\boxed{c^2\,\mathrm{d}\tau^2 \overset{(3.7\text{-}25)}{=} c^2\left(1-\frac{2\,MG}{c^2\,r}\right)\mathrm{d}t^2 - \left(1-\frac{2\,MG}{c^2\,r}\right)^{-1}\mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,sin^2\,\theta\,\mathrm{d}\phi^2} \tag{3.59}$$

$$(g_{\mu\nu}) \overset{(3.59)}{=} \begin{pmatrix} 1-\frac{2\,MG}{c^2\,r} & & & \\ & -\frac{1}{1-\frac{2\,MG}{c^2\,r}} & & \\ & & -r^2 & \\ & & & -r^2\,sin^2\,\theta \end{pmatrix} \tag{3.7-26}$$

where $M$ is the mass of the body creating the gravitational field, and $G$ is the gravitational constant. This finishes the derivation of the Schwarzschild metric tensor.

Note 1 With $M = 0$, equation (3.59) reduces to spherical coordinates in flat spacetime. As $M$ grows, it eventually represents a black hole, and equation (3.59) describes spacetime near a black hole.

Note 2 The result (3.7-26) is a static metric; it does not depend upon $t$. Nothing was required of the body of mass $M$ except that it generate a metric tensor field that is (a) static and (b) spherically symmetric. The body itself need not be static; it could be a collapsing star, as long as the collapse is symmetric.

Note 3 There are no concave-convex “ripples” in “Schwarzschild geometry. The deviation of $g_{\mu\nu}$ from flat spacetime $\eta_{\mu\nu}$ is concave outside a radius of $\frac{2GM}{c^2}$ and convex inside that radius:

$M \geq 0$ and $G > 0$.

$$r > \frac{2GM}{c^2} \Rightarrow g_{\mu\mu} - \eta_{\mu\mu} \overset{(3.7-26)}{\leq} 0 \quad \text{and} \quad r < \frac{2GM}{c^2} \Rightarrow g_{\mu\mu} - \eta_{\mu\mu} \overset{(3.7-26)}{\geq} 0 \qquad \checkmark$$

In Section 4.8, the radius $\frac{2GM}{c^2}$ will be seen to be the Schwarzschild radius for a black hole.

Note 4 When $r \neq \frac{2GM}{c^2}$, we cannot assume that $r$ is radial distance. All we know about $r$ for now is that the surface area of a space sphere $\Sigma$ of radius $r$ (see comment preceding equation 3.52) is $4\pi r^2$. Section 4.1 will explain why we can no longer regard the coordinates $t$ and $r$ as having the same simple meaning as their Euclidean counterparts.

Note 5 The Schwarzsfield solution is NOT a solution to the Lambda empty space equations (3.5-6): $R_{\mu\nu} = \Lambda\, g_{\mu\nu}$. This is because the Schwarzsfield solution is the metric tensor $g_{\mu\nu}$, which means that the calculations above for $R_{\mu\nu}$ remains unchanged in the Lambda scenario. Thus, $R_{\mu\nu} = \Lambda\, g_{\mu\nu}$ holds only if $\Lambda = 0$.

## 3.8 Solution to the Lambda Field Equations (Bud's solution)

We mimic the Schwarzsfield solution approach of the prior section but adopt only the last three assumptions, modifying assumption (c) and (d):

(c) Spacetime is empty *except for dark energy*

(d) Spacetime is asymptotically flat *when there is no dark energy*

(e) Spacetime can be coordinatized by $(t,r,\theta,\phi)$ where **coordinate *t* is timelike**,

Let $c^2\, d\tau^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$ be a solution to the Lambda field equations expressed in spherical coordinates. Set

$$g_{00} = \frac{1}{c^2} A(t,r,\theta,\phi), \quad g_{11} = -B(t,r,\theta,\phi), \quad g_{22} = -r^2, \quad \text{and} \quad g_{33} = -r^2 \sin^2\theta\,. \qquad (3.8\text{-}1)$$

We are not requiring the metric tensor to be static. We allow $A = A(t,r,\theta,\phi)$ to be functions of any or all of the coordinate parameters, and similarly for $B$.

Assumption (c) means that $A$ and $B$ are to be solved using the Lambda vacuum field equations (3.5-6):

$$R_{\mu\nu} = \Lambda\, g_{\mu\nu} \qquad (3.8\text{-}2)$$

Assumption (d) now includes all values of $\Lambda$, including $\Lambda = 0$. So, like Section 3.8, assumption (d) specifies

$$\mathrm{Lim}_{r\to\infty} A \to c^2 \quad \text{and} \quad \mathrm{Lim}_{r\to\infty} B \to 1 \;\; \text{when } \Lambda = 0. \qquad (3.8\text{-}3)$$

The Schwarzschild solution $g_{00} = 1 - \dfrac{2MG}{c^2 r}$ satisfies the empty space equation $R_{00} = 0$, thus it doesn't satisfy the Lambda empty space equation $R_{00} = \Lambda\, g_{00}$ for $\Lambda \neq 0$. We must solve for $A$ and $B$ anew.

The Ricci coefficients from Section 3.8, equations (3.54 – 3.57), still hold because the Ricci formula (3.7-4) for $R_{\mu\nu}$ depends only upon the use of spherical coordinates and the metric tensors along with their partial derivatives with respect to $r$, and these haven't changed. Specifically,

$$R_{00} \overset{(3.55)}{=} -\frac{A'}{B c^2 r} + \frac{(A')^2}{4 A B c^2} + \frac{A' B'}{4 B^2 c^2} - \frac{A''}{2 B c^2} \qquad (3.8\text{-}4)$$

$$R_{11} \overset{(3.55)}{=} -\frac{B'}{B r} - \frac{(A')^2}{4 A^2} - \frac{A' B'}{4 A B} + \frac{A''}{2 A} \qquad (3.8\text{-}5)$$

where prime denotes partial differentiation with respect to $r$.

Since $R_{00}$ and $R_{11}$ are expressed in terms of $A$ and $B$, in order to get $R_{00} = \Lambda\, g_{00}$ and $R_{11} = \Lambda\, g_{11}$, it seems that either $A$ or $B$ (or both) must include $\Lambda$. We do this now, but the implication of doing so is to extend the definition of metric tensor beyond representing just standard mass and energy, now to also include dark energy.

We propose a Schwarzschild-like metric, in the sense that we define $B = \frac{c^2}{A}$. Then $R_{00}$ and $R_{11}$ can each be written in terms of the single variable, $A$:

$$R_{00} = -\frac{A\,(2\,A' + r\,A'')}{2\,c^4\,r} \qquad (3.8\text{-}6)$$

$$R_{11} = \frac{2\,A' + r\,A''}{2\,r\,A} \qquad (3.8\text{-}7)$$

$$R_{22} = -1 + \frac{A + r\,A'}{c^2} \qquad (3.8\text{-}8)$$

Each of equations (3.8-6 to 3.8-8) represent a single equation in a single unknown. Fortunately, the solution to the first two equations is the same:

$$A = -\frac{1}{3}\,c^2 r^2 \Lambda + \frac{k\,c^2}{r} + k_2\,. \qquad (3.8\text{-}9)$$

Unfortunately, this fails to satisfy $\lim\limits_{r\to\infty} A = c^2$, but that is why we chose to modify the meaning of flat spacetime to include the condition that it lacks dark energy. When $\Lambda = 0$, it is indeed true that $\lim\limits_{r\to\infty} A = c^2$ if we set $k_2 = c^2$. Thus, the expression for $A$ becomes

$$A = c^2\left(1 + \frac{k}{r} - \frac{1}{3} r^2 \Lambda\right), \qquad (3.8\text{-}10)$$

and the value of $B$ becomes

$$B = \left(1 + \frac{k}{r} - \frac{1}{3} r^2 \Lambda\right)^{-1}. \qquad (3.8\text{-}11)$$

Solving equation (3.8-8), the third Lambda empty space equation for $A$ directly results in the formula (3.8-10), and this shows that the three equations have the same solution when we choose $c^2$ for the $k_2$ in the first two equations.

We solve for the remaining constant of integration $k$ just as we did in Section 3.7: we compute the equivalent spacetime Euclidean metric tensor $G_{MN}$ and use the fact that $G_{MN} \to \eta_{MN}$ as $r \to \infty$ when $\Lambda = 0$. This can be expressed as

$$G_{00} \to 1, \quad G_{II} \to -1, \quad \text{and} \quad G_{MN} \to 0 \text{ for } M \neq N \quad (\text{when } \Lambda = 0). \qquad (3.8\text{-}12)$$

The transformation equation from $g$ to $G$ is

$$g_{\mu\nu} = \Lambda^M_\mu \, \Lambda^N_\nu \, G_{MN} = \frac{\partial X^M}{\partial x^\mu} \frac{\partial X^N}{\partial x^\nu} G_{MN}. \qquad (3.8\text{-}13)$$

Since $g_{\mu\nu} = 0$ when $\mu \neq \nu$, we are primarily concerned with

$$g_{\mu\mu} = \frac{\partial X^M}{\partial x^\mu} \frac{\partial X^N}{\partial x^\mu} G_{MN}. \qquad (3.8\text{-}14)$$

Recall that $X^0 = ct$, $X^1 = x = r \sin\theta \cos\phi$, $X^2 = y = r \sin\theta \sin\phi$, and $X^3 = z = r\cos\theta$, and $x^0 = ct$, $x^1 = r$, $x^2 = \theta$, and $x^3 = \phi$.

So,

$$g_{00} = \frac{\partial X^M}{\partial x^0} \frac{\partial X^N}{\partial x^0} G_{MN} = \frac{\partial X^M}{\partial (ct)} \frac{\partial X^N}{\partial (ct)} G_{MN} = \left(\frac{\partial (c\,t)}{\partial (c\,t)}\right)^2 G_{00} + 0 + \ldots = G_{00} \qquad (3.8\text{-}15)$$

$$g_{11} = \frac{\partial X^M}{\partial x^1} \frac{\partial X^N}{\partial x^1} G_{MN} = \frac{\partial X^M}{\partial r} \frac{\partial X^N}{\partial r} G_{MN}$$

$$= \left(\frac{\partial x}{\partial r}\right)^2 G_{11} + \left(\frac{\partial y}{\partial r}\right)^2 G_{22} + \left(\frac{\partial z}{\partial r}\right)^2 G_{33} + G_{MN} \text{ terms (where } M \neq N)$$

$$= \sin^2\theta \cos^2\phi \; G_{11} + \sin^2\theta \sin^2\phi \; G_{22} + \cos^2\theta \; G_{33} + G_{MN} \text{ terms} \qquad (3.8\text{-}16)$$

$$\Rightarrow \quad g_{00} \overset{(3.8-15)}{=} G_{00} \overset{(3.8-12)}{\to} 1 \;\checkmark$$

$$\Rightarrow \quad g_{11} \overset{(3.8-16,\, 3.8-12)}{\to} -(\sin^2\theta \cos^2\phi + \sin^2\theta \sin^2\phi + \cos^2\theta) + 0 = -1 \;\checkmark$$

If we express $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, then

$$h_{00} = g_{00} - 1 \overset{(3.8-1)}{=} \frac{A}{c^2} - 1 \overset{(3.8-10)}{=} \frac{k}{r} - \frac{1}{3} r^2 \overset{(3.8-1)}{=} \qquad (3.8\text{-}17)$$

Accepting the quasi-static conditions used to develop equation (2.83), including $h_{00,0} \approx 0$, then $h_{00} \overset{(2.83)}{=} \frac{2V}{c^2}$ holds for any value of $\Lambda$. Since, also, equation (3.8-17) for

$A$ is valid for any constant $\Lambda$, then

$$\frac{k}{r} \overset{(3.8-17)}{=} h_{00} \overset{(2.83)}{=} \frac{2\,V}{c^2} \overset{(2.8-2)}{=} -\frac{2}{c^2}\,\frac{GM}{r}.$$

$$\Rightarrow \quad k = -2\frac{GM}{r}\ . \tag{3.8-18}$$

$$\Rightarrow \quad A = c^2\,(1 - \frac{2\,M\,G}{c^2\,r} - \frac{1}{3}r^2\Lambda) \tag{3.8-19}$$

$$\Rightarrow \quad B = (1 - \frac{2\,M\,G}{c^2\,r} - \frac{1}{3}r^2\Lambda)^{-1} \tag{3.8-20}$$

Using these formulas for $A$ and $B$ generates the following line element and corresponding metric tensor:

$$\begin{aligned} c^2\,d\tau^2 = g_{\mu\nu}\,dx^\mu\,dx^\nu &= A\,dt^2 - B\,dr^2 - r^2\,d\theta^2 - r^2\,sin^2\theta\,d\phi^2 \\ &= c^2\,(1 - \frac{2\,MG}{c^2\,r} - \frac{1}{3}r^2\Lambda)\,dt^2 - (1 - \frac{2\,MG}{c^2\,r} - \frac{1}{3}r^2\Lambda)^{-1}\,dr^2 \\ &\quad - r^2\,d\theta^2 - r^2\,sin^2\theta\,d\phi^2 \end{aligned} \tag{3.8-21}$$

$$g_{\mu\nu} = \begin{pmatrix} 1 - \frac{2\,MG}{c^2\,r} - \frac{1}{3}\,r^2\,\Lambda & & & \\ & -\left(1 - \frac{2\,MG}{c^2\,r} - \frac{1}{3}\,r^2\,\Lambda\right)^{-1} & & \\ & & -r^2 & \\ & & & -r^2\,sin^2\,\theta \end{pmatrix} \tag{3.8-22}$$

This constitutes an exact solution for Einstein’s Lambda field equations.

# Chapter 4 Physics in the vicinity of a massive object

## 4.0 Introduction

The Schwarzschild solution provides the line element for measuring distances between points that are exterior to some object with mass. The metric tensor in the Schwarzschild solution is a function of the mass. This allows us to explore spacetime from flat (where $m = 0$) all the way to the extreme curvature that occurs just outside a black hole (where $m$ is large).

Section 4.1 investigates the physical meaning of coordinates like $t$, $r$, $\theta$, and $\phi$ in curved spacetime.

Sections 4.2 and 4.3 develop equations of physics for things like the parallax of Venus that enable general relativity to make predictions.

Sections 4.4 – 4.6 explore the geodesics of light and other massless particles.

Section 4.7 develops parallel transport in spacetime for particles and includes the geodesic effect that that due to curvature, parallel transport around a circle can now experience a change in the tangent angle.

Section 4.8 develops Eddington-Finkelstein coordinates, an alternate Schwarzschild coordinate system that is valid inside the Schwarzschild radius.

Section 4.9 develops isotropic coordinates, an additional Schwarzschild coordinate system that reduces to Cartesian coordinates rather than spherical coordinates. It also develops Kruskal-Szekeres coordinates in which null geodesics are lines sloped at 45°.

Section 4.10 develops the Kerr solution, a generalization of the Schwarzschild solution, that applies to rotating black holes in otherwise empty spacetime. It shows that for rotating black holes there are two event horizons, not one.

## 4.1 Length and time

Setting $m = \frac{GM}{c^2}$ allows us explore Schwarzschild's solution using more compact notation:

$$c^2\,\mathrm{d}\tau^2 \overset{(3.59)}{=} c^2\left(1 - \frac{2m}{r}\right)\mathrm{d}t^2 - \frac{1}{1-\frac{2m}{r}}\,\mathrm{d}r^2 - r^2\,\mathrm{d}\theta^2 - r^2\,\sin^2\theta\,\mathrm{d}\phi^2 \tag{4.1}$$

where $r > 0$, $m \geq 0$.

We note that $m$ has units of distance, important because the Schwarzschild radius will be define in Section 4.8 to be $2m$.

The line element (4.1) is an expression in terms of the metric tensor:

$$c^2\,\mathrm{d}\tau^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \tag{4.1-1}$$

where

$$g_{00} = \left(1 - \frac{2m}{r}\right),\quad g_{11} = -\frac{1}{1-\frac{2m}{r}},\quad g_{22} = -r^2,\ \text{ and }\ g_{33} = -r^2\,\sin^2\theta \tag{4.1-2}$$

and

$$x^0 = ct,\ \ x^1 = r,\ \ x^2 = \theta,\ \text{ and }\ x^3 = \phi. \tag{4.1-3}$$

Define a metric tensor $\tilde{g}_{\mu\nu}$ for the 3-dimensional manifold S that results from freezing $t$:

$$\tilde{g}_{ij} \equiv -g_{ij}. \tag{4.1-4}$$

The line element for S is

$$\mathrm{d}s^2 = \tilde{g}_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j = \frac{1}{1-\frac{2m}{r}}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2 + r^2\,\sin^2\theta\,\mathrm{d}\phi^2. \tag{4.2}$$

$\tilde{g}_{ij}$ is a positive definite metric tensor:

Positive definite means that

$$\tilde{g}_{ij}x^i x^j \geq 0\ \forall\, i, j\ \ \text{and}\ \ \tilde{g}_{ij}x^i x^j = 0 \text{ only if } x^i = 0 \text{ for all } i.$$

First, $\tilde{g}_{ii} > 0$ for $i = 1\text{–}3$:

$$1 - \frac{2m}{r} \overset{(3.7\text{–}24)}{>} 0 \quad\Rightarrow\quad g_{11} \overset{(4.1\text{–}2)}{<} 0 \quad\Rightarrow\quad \tilde{g}_{11} \overset{(4.1\text{–}4)}{>} 0.$$

Since $\tilde{g}_{ij}$ is not defined on the $z$-axis or at the origin, we also have

$\tilde{g}_{22} = r^2 > 0$ and $\tilde{g}_{33} = r^2 sin^2\theta > 0$ . ✓

Next, $x^1 = r > 0$, $x^2 = \theta > 0$, and $x^3 = \phi \geq 0$.

Thus, $\tilde{g}_{ij}x^i x^j > 0$:

$\tilde{g}_{ij}x^i x^j = \tilde{g}_{11}x^1 x^1 + \tilde{g}_{22}x^2 x^2 + \tilde{g}_{33}x^3 x^3$

The third term is nonnegative but each of the first two terms is positive. ✓

This makes $\tilde{g}_{ij}$ positive definite. ✓

$\tilde{g}_{ij}$ being positive definite means that the 3-dimensional manifold S is a slice of space rather than spacetime. Moreover, none of the $\tilde{g}_{ij}$ depend on $t$, so we can refer to events with the same $(r,\theta,\phi)$ but different $t$ as occurring at the same point in space. We call these **fixed points of space**, and this process of separating spacetime into space and time is called a **spacetime split.** This is not a general feature of spacetime. However, when it happens, it is much easier to solve problems because we are working in static space rather than in non-static spacetime.

When $M = 0$, the line element (4.1) becomes that of flat spacetime in spherical coordinates, while the line element (4.2) becomes that of Euclidean space in spherical coordinates. As $M$ grows, it causes increasing distortion in both spacetime and space so that neither is flat. Distortion is measured by the dimensionless quantity $\frac{2m}{r}$ and is maximized at the boundary of the object, the smallest the radius $r$ can achieve.

The table below shows how much distortion is caused by objects of various radius, $r_B$, and mass, $M$. Distortion around our sun is small enough to ignore for most calculations but is large enough to show up in precision measurements of the perihelions of Mercury, Venus, and Earth and also in the deflection of light. Treating the distortion is necessary for white dwarfs, and of central relevance for a neutron star. The Schwarzschild equations must be modified in order to be relevant for black holes.

Table 4.1-1

| Object | $r_B$ (meters) | $M$ (kg) | $2m$ | $2m / r_B$ | Distortion |
|---|---|---|---|---|---|
| Flat spacetime | - | - | - | 0 | None |
| Proton | $8\text{x}10^{-16}$ | $2\text{x}10^{-27}$ | $2\text{x}10^{-54}$ | $10^{-39}$ | Infinitessimal |
| Earth | $6\text{x}10^{6}$ | $6\text{x}10^{24}$ | $9\text{x}10^{-3}$ | $10^{-9}$ | Very tiny |
| Sun | $7\text{x}10^{8}$ | $2\text{x}10^{30}$ | $3\text{x}10^{4}$ | $10^{-6}$ | Tiny |
| White Dwarf | $7\text{x}10^{6}$ | $2\text{x}10^{30}$ | $2\text{x}10^{3}$ | $10^{-4}$ | Important |
| Neutron Star | $1\text{x}10^{4}$ | $4\text{x}10^{30}$ | $6\text{x}10^{3}$ | .6 | Huge |
| Black Hole | 0 | Up to $10^{40}$ | Up to $10^{13}$ | Undefined | Undefined |

In flat spacetime, $r$ is simply the distance from the origin, but for $M > 0$, $r$ has a positive lower bound, the radius $r_B$, and the meaning of $r$ is not so clear. What does it represent?

For $M > 0$, the space manifold S is curved. We consider two different submanifolds of S, a sphere and a line. Let Σ be a sphere in space having radius $r$. Since d$r$ = 0 on Σ, equation (4.2) reduces to the same equation as the sphere Σ from Section 3.7 whose line element is

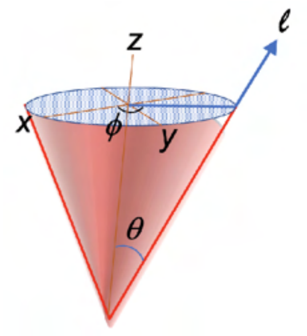
$$\mathrm{d}L^2 \overset{(3.52)}{=} r^2 \,(\mathrm{d}\theta^2 + sin^2\theta\, \mathrm{d}\phi^2). \tag{4.3}$$

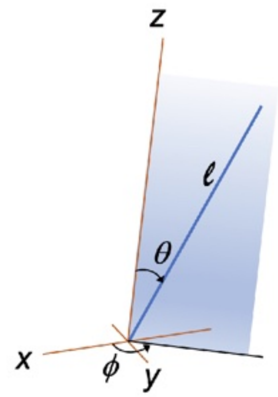
It follows that Σ has the 2-dimensional geometry of a sphere of radius $r$ in flat Euclidean space, and, so, infinitesimal (tangential) distances between points on S are given by

$$\mathrm{d}L = r\sqrt{\mathrm{d}\theta^2 + sin^2\theta\, \mathrm{d}\phi^2}\,. \tag{4.4}$$

That is, the distance between points is the familiar Euclidean distance.

However... let $\ell$ be the radial line generated by constant $\phi$ and constant $\theta$. See the first figure below where $\ell$ is shown as the line at angle $\phi$ in the cone of angle $\theta$, or the second figure that shows $\ell$ as the line at angle $\theta$ in the plane of constant $\phi$, and also in the figure in Example 1.1.4.




The line element for ℓ can be obtained from equation (4.2) by setting $d\theta = d\phi = 0$:

$$dR^2 = \frac{1}{1-\frac{2m}{r}}\, dr^2$$

and, so, the infinitesimal (tangential) distance between points is

$$\boxed{dR = \frac{1}{\sqrt{1-\frac{2m}{r}}}\, dr}. \tag{4.5}$$

Distance between points is obtained by integrating $dR$ over a geodesic between the points. That is, $R$ measures the distance between points. Because $dR > dr$, $r$ no longer measures coordinate distance.

We have just learned that length can be Euclidean in one submanifold (the sphere) yet be non-Euclidean in another submanifold (the line) of the same space S. This conflicting story is explained by curvature.

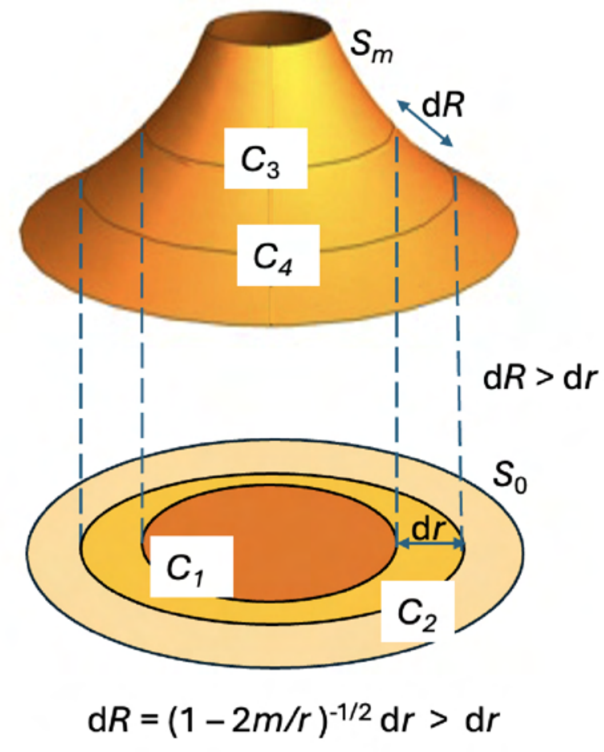



Figure 4.1 Radial Distance in the Schwarzschild Geometry

Figure 4.1 shows Euclidean spheres of radii $r$ and $r + \mathrm{d}r$ embedded in both a flat space $S_0$ and a curved space $S_m$. The distance between the spheres is d$r$ in flat Euclidean space but is d$R$ > d$r$ in curved space, where d$R$ is given in equation (4.5). The range of $r$ is $(2m, \infty)$ and we see that $\lim_{r\to\infty} \mathrm{d}R = \mathrm{d}r$ and $\lim_{r\to 2m} \mathrm{d}R = \infty$. For a stellar object like a planet or a star, $r$ cannot reach the lower limit $2m$ because it reaches $r_B$, the radius of the object, first. That is, $r > r_B$ and $r_B > 2m$.

We conclude that in the manifold S, $r$ can measure length in some cases but be smaller than length in some other cases, distorted by a factor of $\frac{1}{\sqrt{1-\frac{2m}{r}}}$.

Switching topics now to time, recall that in special relativity, clocks record proper time $\tau$ along their world lines. For a stationary clock at location $(r,\theta,\phi)$, we can find the relation between $t$ *and* $\tau$ by setting $\mathrm{d}r = \mathrm{d}\theta = \mathrm{d}\phi = 0$ in equation (4.1):

$$\mathrm{d}\tau = \sqrt{1 - \frac{2m}{r}}\ \mathrm{d}t. \qquad (4.6)$$

Thus, in flat spacetime, where $m = 0$, $d\tau = dt$ records coordinate time. However, in curved spacetime, where $m > 0$, a stationary clock's measurement of coordinate time is distorted by a factor of $\sqrt{1 - \frac{2m}{r}}$. Notice that as $r \to \infty$, $d\tau \to dt$, and the distortion disappears.

## 4.8 Black Holes

The Schwarzschild solution uses coordinates ($t$,$r$,$\theta$,$\phi$). The metric tensor $g_{\mu\nu}$ does not depend on $t$, so the solution is static and we can assume $t$ is a fixed finite number, possibly negative. Reviewing Table 4.1-1, we see that the value of $r$ can decrease from infinity until it reaches either the boundary of the object, $r_B$, or the lower limit of 2$m$, where $m = \frac{\text{GM}}{c^2}$.

Definition An object for which $r_B \leq 2m$ is called a **black hole,** and the distance $\frac{2\,\text{GM}}{c^2}$ is known as the **Schwarzschild radius**.

Recall the Schwarzschild line element, equation (4.1):

$$c^2\, \mathrm{d}\tau^2 \overset{(3.59)}{=} c^2\, (1 - \frac{2\,m}{r})\, \mathrm{d}t^2 - \frac{1}{1-\frac{2\,m}{r}}\, \mathrm{d}r^2 - r^2\, \mathrm{d}\theta^2 - r^2\, sin^2\theta\, \mathrm{d}\phi^2$$

We see that $g_{11} \to -\infty$ as $r \to 2m$. However, metric tensor singularities are coordinate-dependent. We might refer to these as “coordinate singularities”. For example, in Euclidean 3-space $\mathbb{E}^3$ we saw that the spherical coordinates dual basis line element exhibits singularities in the metric tensor at $r$ = 0 and also at $\phi$ = 0. But, these same points have finite metric tensor values in other coordinate systems like Euclidean coordinates (which has no singularities). We develop alternate Schwarzschild geometry coordinate systems below in which the metric tensor has a finite value at $r$ = 2m, showing it is not a real singularity. We define a **singularity** as a coordinate singularity that holds for all coordinate systems. (This can be difficult to show. For example, the Schwarzschild singularity at $r$ = 0 is a real singularity.)

Because of the coordinate singularity, Schwarzschild coordinates ($t$,$r$,$\theta$,$\phi$) are inadequate for treating $r \leq 2m$ . To explore null trajectories (i.e., photons) near the lower limit 2$m$, various alternate coordinate systems for Schwarzschild geometry have been developed that do not have coordinate singularities at $r$ = 2$m$, including Eddington-Finkelstein coordinates (the actual equations were developed by Roger Penrose), Kruskai-Szekeres coordinates, Lemaitre coordinates, and Gullistrand-Painleve coordinates. Vaidya has developed a generalized version of the Schwarzschild solution by extending the Schwarzschild line element to allow the mass $M$ to vary as a function of the coordinates.

We develop the Eddington-Finkelstein coordinate systems. There are two coordinate systems, Out-going (“retarded”) and In-going (“Advanced”). In the In-going coordinate system, we replace $t$ by

$$v = ct + r + 2m \ln\left|\frac{r}{2m} - 1\right|. \tag{4.60}$$

This system allows us to watch a photon fall into a black hole. The absolute value sign handles the situation when $r_B < r < 2m$.

In the Out-going coordinate system, we replace $t$ by

$$u = ct - r - 2m \ln\left|\frac{r}{2m} - 1\right|.$$

This system allows us to analyze null trajectories (photons) inside a black hole.

The Schwarzschild line element (4.1) for the In-going coordinate system becomes (Exercise 4.8.1)

$$c^2\, d\tau^2 = \left(1 - \frac{2m}{r}\right) dv^2 - 2\, dv\, dr - r^2\, d\theta^2 - r^2 \sin^2\theta\, d\phi^2 \tag{4.61}$$

Notice that the coordinate singularity as $r \to 2m$ has disappeared.

Definition $(v, r, \theta, \phi)$, where $r > r_B$, are called **Eddington-Finkelstein coordinates**.

These coordinates are defined for all $v$, $\theta$, and $\phi$, and for $r > r_B$.

Since $c^2\,(d\tau)^2 \equiv g_{\mu\nu}\, dx^\mu\, dx^\nu$ where $x^0 = v$, the metric tensor is

$$g_{\mu\nu} = \begin{pmatrix} 1 - \frac{2m}{r} & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$

When $r = 2m$, none of the metric tensor components becomes infinite, so unlike Schwarzschild coordinates, we can analyze photons that cross the Schwarzschild radius by setting $r_B < 2m$.

Photon trajectories are null geodesics ($d\tau = 0$). The path of a photon falling radially inward is the line of intersection of the cone, $\theta$ = constant, and the plane, $\phi$ = constant.

So, $\frac{d\theta}{dr} = \frac{d\phi}{dr} = 0$, and the null line element can be expressed as

$$0 = c^2 \left(\frac{d\tau}{dr}\right)^2 \overset{(4.61)}{=} (1 - \frac{2\,m}{r}) \left(\frac{dv}{dr}\right)^2 - 2\,\frac{dv}{dr} - r^2 \left(\frac{d\theta}{dr}\right)^2 - r^2 \sin^2\theta \left(\frac{d\phi}{dr}\right)^2$$

$$= \frac{dv}{dr}\,[\,(1 - \frac{2\,m}{r})\,\frac{dv}{dr} - 2\,]$$

$$\Rightarrow \quad \frac{dv}{dr} = 0 \qquad (4.8\text{-}1)$$

or

$$\frac{dv}{dr} = \frac{2}{1 - \frac{2\,m}{r}} \qquad (4.8\text{-}2)$$

$$\frac{dv}{dr} \overset{(4.60)}{=} c\,\frac{dt}{dr} + 1 + \frac{2\,m\,\frac{1}{2\,m}}{\frac{r}{2\,m} - 1} = c\,\frac{dt}{dr} + 1 + \frac{2\,m}{r-2\,m} = c\,\frac{dt}{dr} + \frac{r-2\,m+2\,m}{r-2\,m}$$

$$= \ c\,\frac{dt}{dr} + \frac{r}{r-2\,m} = \ c\,\frac{dt}{dr} + \frac{1}{1 - \frac{2\,m}{r}} \qquad (4.8\text{-}3)$$

$$\frac{dv}{dr} = 0$$

$\overset{(4.8-3)}{\Rightarrow} \quad c\,\frac{dt}{dr} = -\,\frac{1}{1 - \frac{2\,m}{r}} < 0$ if $r > 2m$ (i.e., for an in-going photon outside of $2m$)

$$\frac{dv}{dr} = \frac{2}{1 - \frac{2\,m}{r}}$$

$$\overset{(4.8-3)}{\Rightarrow} c\,\frac{dt}{dr} + \frac{1}{1 - \frac{2\,m}{r}} = \frac{2}{1 - \frac{2\,m}{r}} = c\,\frac{dt}{dr} + \frac{1}{1 - \frac{2\,m}{r}} > 0$$

if $r > 2m$ (i.e., for an out-going photon outside of $2m$)

Integrating equation (4.8-1) yields

$v = A$ (where A is a constant) (In-going) (4.8-4)

Observe that equation (4.8-2) can be written $\frac{dv}{dr} = 2 + \frac{4\,m}{r-2\,m}$:

$$\frac{dv}{dr} = \frac{2}{1-\frac{2m}{r}} = \frac{2r}{r-2m} = \frac{2r-4m}{r-2m} + \frac{4m}{r-2m} = 2 + \frac{4m}{r-2m} \quad ✓$$

So, integrating equation (4.8-2) yields

$v = 2r + 4m \ln|r - 2m| + B$ (where $B$ is a constant) (Out-going) (4.8-5)|

We observe that for radial null rays (4.8-1) and (4.8-2), as $r \to \infty$, dv/dr → 0 and ±2, respectively, not ±1 as one might expect if one regarded $v$ as "time".

Equations (4.8-4) and (4.8-5) enable us to draw spacetime plots having axes $v$ and $r$. Just as we did in flat spacetime, we draw null radials $v = A$ at 45°. This includes the $r$-axis, which is $v = 0$. Figure 4.1.3, to the right of the vertical line $r = 2m$, shows the in-going and out-going light rays. The out-going null radials approach -∞ at $r = 2m$.

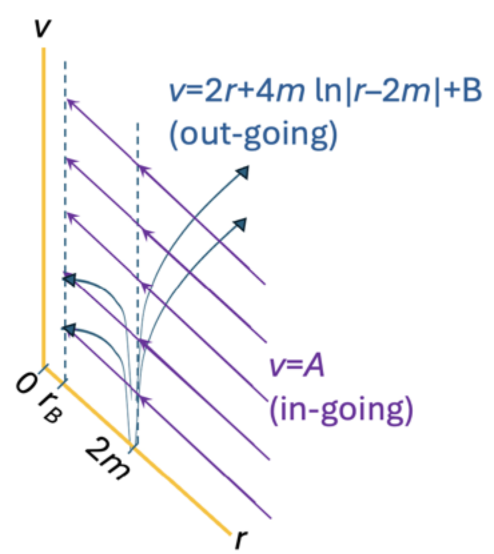



Figure 4.1.3 Eddington-Finkelstein In-going and Out-going Null Geodesics

In addition, equations (4.8-4) and (4.8-5) can be valid to the left of the vertical line $r = 2m$ by letting $r_B < 2m$ since both equations are well-defined for $0 \le r_B < r < 2m$. In both cases, $\frac{dv}{dr} < 0$ when $r < 2m$, so Figure 4.8-1 appropriately plots both in-going and out-going trajectories as in-going. This illustrates what is meant when one says that light cannot escape from inside a black hole. Because light cannot escape from a sphere of radius $2m$, an outside observer cannot see such events, and we make the following definition.

Definition The hyperspace “sphere” of “radius” $2m$ in spacetime is called an **Event Horizon**.