11

Tensors

Actually, we've been doing tensor analysis all along. All we will do now is to add a few elements which will make things look even more like "classical tensor analysis".

In all the following, we assume that we are dealing with position in some N-dimensional space, and that

$$\{(x^1, x^2, \ldots, x^N)\} \quad , \quad \{h_1, h_2, \ldots, h_N\} \quad \text{and} \quad \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\} \quad .$$

is some coordinate system with associated scaling factors and unit tangent vectors. We will not automatically assume this is an orthogonal system.

11.1 Two Standard Conventions

The Kronecker Delta

One element of the "classical tensor analysis look" is a slight extension of the Kronecker delta notation. For any pair of integers i and j, we'll let

$$\delta^i{}_j = \delta_i{}^j = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases} \quad .$$

We'll treat these as constant scalar fields on our space.

The Einstein Convention

Einstein used a shorthand for summations that many others dealing with tensor analysis decided to copy. Unfortunately, there are at least two versions of the "Einstein convention":

Version 1: If an expression involves the product of two indexed quantities sharing one particular index, then you should assume that these products are summed up over that index. Thus,

$$A_{ij} B_j \qquad \text{means} \qquad \sum_{j=1}^{N} A_{ij} B_j$$

where $j = 1, 2, \ldots, N$.

Version 2: If an expression involves the product of two indexed quantities sharing one particular index *with one having the index as a superscript and the other having the index as a subscript*, then you should assume that these products are summed up over that index. Thus,

$$A_{ij} B^j \qquad \text{means} \qquad \sum_{j=1}^{N} A_{ij} B^j$$

where $j = 1, 2, \ldots, N$. Often (but not always) you do NOT sum up over a repeated index when both are superscripts or both are subscripts (in classical tensor analysis, repeated subscripts or repeated superscripts rarely happen).

Either version of the Einstein convention can be convenient in private computations, especially in those involving very many summations. However, neither would be of appreciable value in our discussions, and may even cause some moments of confusion to those first encountering it. So we will keep the summation symbols, and NOT use any version of the Einstein convention.

11.2 The Reciprocal Basis Fields

Remember that, at each point, $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}$ is a basis for the tangent vector space at that point. There are two other bases that are commonly used for each tangent space,

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \qquad \text{and} \qquad \left\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\right\} \quad .$$

The first, we've already seen. For each k ,

$$\boldsymbol{\varepsilon}_k = \frac{\partial \boldsymbol{r}}{\partial x^k} = h_k \mathbf{e}_k \quad .$$

We often used $\mathbf{e}_k$'s instead of the $\boldsymbol{\varepsilon}_k$'s simply because we normally prefer unit basis vectors.

The other basis, $\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\}$, is the basis reciprocal to $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\}$. Recalling the definition of reciprocal bases from section 2.6, we recall that each $\boldsymbol{\varepsilon}^i$ at each point in space is chosen to be the single vector such that:

1. $\boldsymbol{\varepsilon}^i$ is orthogonal to all the $\boldsymbol{\varepsilon}_j$'s except $\boldsymbol{\varepsilon}_i$ (hence $\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}_j = 0$ if $i \neq j$).
2. The scalar projection of $\boldsymbol{\varepsilon}^i$ onto $\boldsymbol{\varepsilon}_i$ is $1/\|\boldsymbol{\varepsilon}_i\|$ (hence $\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}_i = 1$).

More concisely,

$$\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}_j = \delta^i{}_j \quad . \tag{11.1}$$

Keep in mind that neither basis need be orthogonal. So the above does not (necessarily) mean that each $\boldsymbol{\varepsilon}^i$ is parallel to $\boldsymbol{\varepsilon}_i$ (see figure 11.1).

?►*Exercise 11.1:* *Show that,* if the coordinate system is orthogonal, *then*

$$\boldsymbol{\varepsilon}^i = \frac{1}{h_i} \mathbf{e}_i \quad .$$

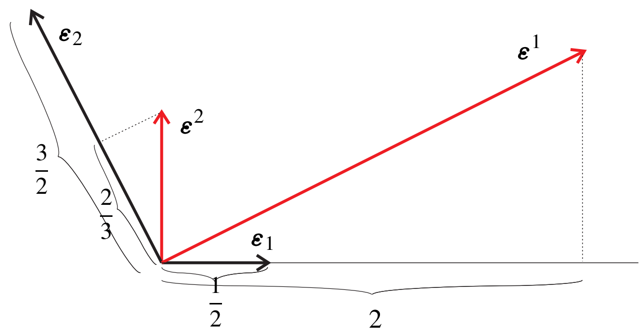

Figure 11.1: A basis for a two-dimensional vector space (in black) with $\|\boldsymbol{\varepsilon}_1\| = 1/2$ and $\|\boldsymbol{\varepsilon}_2\| = 3/2$, and the corresponding reciprocal basis (in red).

11.3 Co- and Contravariant Components of Vector Fields

Definitions

Now let $\mathbf{F}$ be any vector field. At each point, it can be expressed in terms of either of these bases,

$$\mathbf{F} = \sum_{i=1}^{N} F^i \boldsymbol{\varepsilon}_i \qquad \text{or} \qquad \mathbf{F} = \sum_{j=1}^{N} F_j \boldsymbol{\varepsilon}^j \quad .$$

The F^i's (i.e., the components with respect to the $\boldsymbol{\varepsilon}_i$'s) are called the *contravariant* components of $\mathbf{F}$ (with respect to the given coordinate system), and the F_i's (i.e., the components with respect to the $\boldsymbol{\varepsilon}^i$'s) are called the *covariant* components of $\mathbf{F}$ (with respect to the given coordinate system).[1]

It is important that you (re)do the next exercise (preferably without looking back at section 2.6). It explains why using reciprocal bases is such a clever thing to do.

?►*Exercise 11.2:* *Let* $\mathbf{V}$ *and* $\mathbf{U}$ *be two vector fields. Using the co- and contravariant components of each, show that*

$$\mathbf{V} \cdot \mathbf{U} = \sum_{i=1}^{N} V^i U_i = \sum_{i=1}^{N} V_i U^i \quad . \tag{11.2}$$

As a not-very-significant corollary of this exercise, we immediately have the following:

Corollary 11.1 (Quotient Rule for vector fields)
Suppose ψ *is a scalar field on a region* $\mathcal{R}$, *and we have two sets of* N *other scalar fields*

$$\left\{ V^1, V^2, \ldots, V^N \right\} \qquad \text{and} \qquad \{ U_1, U_2, \ldots, U_N \}$$

[1] Notice that we are superscripting both the *co*ordinates and the *contra*variant components of vector fields. This will not help keep terminology straight.

such that

$$\sum_{i=1}^{N} V^i U_i = \psi \qquad \text{everywhere in} \quad \mathcal{R} \quad .$$

Then

$$\mathbf{V} = \sum_{i=1}^{N} V^i \boldsymbol{\varepsilon}_i \qquad \text{and} \qquad \mathbf{U} = \sum_{i=1}^{N} U_i \boldsymbol{\varepsilon}^i$$

are vector fields on $\mathcal{R}$ *with*

$$\mathbf{V} \cdot \mathbf{U} = \psi \quad .$$

A more general "quotient rule" is mentioned in Arfken, Weber and Harris[2] asserting the following

> *Let* $\mathbf{V}$ *and* ψ *be, respectively, a vector field and a scalar field, and suppose we have a set of* N *general formulas, applicable no matter what coordinate system we have, that define* N *scalar fields* U_1 , U_2 , … *and* U_N *(these* U_k*'s will change with different coordinate systems). Suppose further that, no matter what coordinate system we have, the corresponding* U_k*'s satisfy*
>
> $$\sum_{k=1}^{N} V^k U_k = \psi$$
>
> *where the* V^k*'s are the contravariant components of* $\mathbf{V}$ *with respect to the coordinate system. Then there is a single vector field* $\mathbf{U}$ *such that*
>
> $$\mathbf{U} = \sum_{i=1}^{N} U_i \boldsymbol{\varepsilon}^i$$
>
> *in each coordinate system.*

Without additional conditions on how those N general formulas defining the U_i's change as coordinate systems are changed, this "quotient rule" cannot be accepted as truly valid.[3] Instead, as stated, this "quotient rule" is more of a strong suggestion that such a single vector field $\mathbf{U}$ exists.

Some Examples

!►*Example 11.1 (Contravariant components of velocity):* *Let*

$$\boldsymbol{r}(t) \sim \left(x^1(t),\, x^2(t),\, \ldots,\, x^N(t) \right)$$

be some (differentiable) position-valued function (a parametrization for some curve C *). Recall that*

$$\frac{d\boldsymbol{r}}{dt} = \sum_{i=1}^{N} \frac{\partial \boldsymbol{r}}{\partial x^i} \frac{dx^i}{dt} = \sum_{i=1}^{N} \underbrace{h_i \mathbf{e}_i}_{\boldsymbol{\varepsilon}_i} \frac{dx^i}{dt} \quad .$$

---

[2] Page 211 — and they only refer to rotated Cartesian systems

[3] After all, for given vector and scalar fields $\mathbf{V}$ and ψ , the *one* equation $\sum_{k=1}^{N} V^k U_k = \psi$ has the N unknowns U_1 , … , and U_N in each coordinate system. There are infinitly many possible solutions in each coordinate system!

So

$$\frac{d\boldsymbol{r}}{dt} = \sum_{i=1}^{N} \frac{dx^i}{dt} \boldsymbol{\varepsilon}_i$$

which means that ${}^{dx^i}/_{dt}$ *is the* i^{th} *contravariant component of the vector field* ${}^{d\boldsymbol{r}}/_{dt}$ *on the curve* $\mathcal{C}$.

!►*Example 11.2 (Covariant components of the gradient):* *Let* Ψ *be a scalar field with coordinate formula* ψ ,

$$\Psi(\boldsymbol{r}) = \psi\left(x^1, x^2, \ldots, x^N\right) \qquad \text{where} \quad \boldsymbol{r} \sim \left(x^1, x^2, \ldots, x^N\right) \quad .$$

Recall that the most general definition of the gradient of Ψ , $\nabla\Psi$, *is that it is the vector field such that*

$$\frac{d}{dt}\left[\Psi\left(\boldsymbol{r}(t)\right)\right] = \nabla\Psi \cdot \frac{d\boldsymbol{r}}{dt}$$

for any differentiable position-valued function $\boldsymbol{r}(t)$. *Recall, also that we found the formula for* $\nabla\Psi$ *to be*

$$\nabla\Psi = \sum_{k=1}^{N} \frac{1}{h_k} \frac{\partial\psi}{\partial x^k} \mathbf{e}_k$$

provided *the coordinate system is orthogonal. From exercise 11.1, we know* $\boldsymbol{\varepsilon}^k = {}^1/_{h_k}\, \mathbf{e}_k$ *when the coordinate system is orthogonal, as so the above formula can be rewritten as*

$$\nabla\Psi = \sum_{k=1}^{N} \frac{\partial\psi}{\partial x^k} \boldsymbol{\varepsilon}^k$$

provided *the coordinate system is orthogonal. Thus, if the coordinate system is orthogonal, then* ${}^{\partial\psi}/_{\partial x^k}$ *is the* k^{th} *covariant coordinate of* $\nabla\Psi$. *Could*

$$\nabla\Psi = \sum_{k=1}^{N} \frac{\partial\psi}{\partial x^k} \boldsymbol{\varepsilon}^k$$

be the general formula for the gradient in any coordinate system, orthogonal or not? Well, if we assume this formula and let

$$\boldsymbol{r}(t) \sim \left(x^1(t),\, x^2(t),\, \ldots,\, x^N(t)\right)$$

be any (differentiable) position-valued function (as in the previous example), then by the classical chain rule and the results from exercise 11.2 (specifically, formula (11.2) for the dot product), we have

$$\begin{aligned}
\frac{d}{dt}\left[\Psi\left(\boldsymbol{r}(t)\right)\right] &= \frac{d}{dt}\left[\psi\left(x^1(t),\, x^2(t),\, \ldots,\, x^N(t)\right)\right] \\
&= \sum_{i=1}^{N} \frac{\partial\psi}{\partial x^i} \frac{dx^i}{dt} \\
&= \left(\sum_{k=1}^{N} \frac{\partial\psi}{\partial x^k} \boldsymbol{\varepsilon}^k\right) \cdot \left(\sum_{i=1}^{N} \frac{dx^i}{dt} \boldsymbol{\varepsilon}_i\right) = \nabla\Psi \cdot \frac{d\boldsymbol{r}}{dt} \quad ,
\end{aligned}$$

just as we should have. This strongly suggests that, indeed, the general (covariant) formula for the gradient is

$$\nabla\Psi = \sum_{k=1}^{N} \frac{\partial\psi}{\partial x^k}\boldsymbol{\varepsilon}^k \quad .$$

However, have not yet confirmed that this formula defines a vector field independent of the choice of coordinates. A real cynic may suggest the possibility that this formula with two different nonorthogonal coordinate systems could lead to two different vector fields that happen to satisfy the above. This time, that cynic would be wrong — the $^{\partial\psi}/_{\partial x^i}$*'s are the covariant components of* $\nabla\Psi$ *. We'll just have to develop a little more theory to confirm that the above really nice formula is, indeed, coordinate independent.*

11.4 Converting Between Co- and Contravariant Representations

Our goal is now is to determine how to (easily) find the covariant components of a vector field from its contravariant components, and its contravariant components from its covariant components. We start by finding convenient formulas for expressing the vectors in either one of the bases

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \qquad \text{or} \qquad \left\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\right\}$$

in terms of the other.

Components of the Reciprocal Bases with Respect to Each Other

Keep in mind that, at each point,

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \qquad \text{and} \qquad \left\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\right\}$$

are both bases for the same tangent space of vectors at that point. So each $\boldsymbol{\varepsilon}_i$ can be written in terms of the $\boldsymbol{\varepsilon}^j$'s, and each $\boldsymbol{\varepsilon}^j$ can be written in terms of the $\boldsymbol{\varepsilon}_i$'s,

$$\boldsymbol{\varepsilon}_i = \sum_{k=1}^{N} \alpha_{ik}\,\boldsymbol{\varepsilon}^k \qquad \text{and} \qquad \boldsymbol{\varepsilon}^j = \sum_{k=1}^{N} \beta^{jk}\,\boldsymbol{\varepsilon}_k \quad .$$

Now, also recall the relations between the dot products of these basis vectors and the metric and the Kronecker delta,

$$\boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}_j = g_{ij} \qquad \text{and} \qquad \boldsymbol{\varepsilon}^k \cdot \boldsymbol{\varepsilon}_j = \delta^k{}_j \quad .$$

(The first was actually the definition of the [covariant] components of the metric [see page 8–29], and the second was essentially the defining formula for the reciprocal basis.) Combining all the above, we get

$$g_{ij} = \boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}_j = \left(\sum_{k=1}^{N} \alpha_{ik}\,\boldsymbol{\varepsilon}^k\right) \cdot \boldsymbol{\varepsilon}_j = \sum_{k=1}^{N} \alpha_{ik}\,\boldsymbol{\varepsilon}^k \cdot \boldsymbol{\varepsilon}_j = \sum_{k=1}^{N} \alpha_{ik}\,\delta^k{}_j = \alpha_{ij} \quad .$$

So

$$\boldsymbol{\varepsilon}_i = \sum_{j=1}^{N} g_{ij}\,\boldsymbol{\varepsilon}^j \qquad \text{for} \quad i = 1, 2, \ldots, N \quad . \tag{11.3}$$

In other words, the covariant components of the metric are also the covariant components of the $\boldsymbol{\varepsilon}_i$'s.

To get the contravariant components of the $\boldsymbol{\varepsilon}^j$'s (the β^{jk}'s), we go back to elementary matrix theory and rewrite the above formula for the $\boldsymbol{\varepsilon}_i$'s in matrix form,

$$\begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix} = \mathbf{G} \begin{bmatrix} \boldsymbol{\varepsilon}^1 \\ \boldsymbol{\varepsilon}^2 \\ \vdots \\ \boldsymbol{\varepsilon}^N \end{bmatrix} \qquad \text{where} \quad [\mathbf{G}]_{ij} = g_{ij} \quad .$$

Then, of course,

$$\begin{bmatrix} \boldsymbol{\varepsilon}^1 \\ \boldsymbol{\varepsilon}^2 \\ \vdots \\ \boldsymbol{\varepsilon}^N \end{bmatrix} = \mathbf{G}^{-1} \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix} \quad .$$

Because of this relation, it is natural to define the *contravariant components of the metric* (with respect to the given coordinate system) as the corresponding entries in $\mathbf{G}^{-1}$, and to use g^{ij} to denote these quantities. That is,

$$g^{ij} = \left[\mathbf{G}^{-1}\right]_{ij} \qquad \text{where} \quad [\mathbf{G}]_{mn} = g_{mn} \quad .$$

By the above, we then have

$$\boldsymbol{\varepsilon}^i = \sum_{k=1}^{N} g^{ik}\,\boldsymbol{\varepsilon}_k \qquad \text{for} \quad i = 1, 2, \ldots, N \quad . \tag{11.4}$$

From the last formula, we see that

$$\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}^j = \left(\sum_{k=1}^{N} g^{ik}\,\boldsymbol{\varepsilon}_k\right) \cdot \boldsymbol{\varepsilon}^j = \sum_{k=1}^{N} g^{ik}\,\boldsymbol{\varepsilon}_k \cdot \boldsymbol{\varepsilon}^j = \sum_{k=1}^{N} g^{ik}\,\delta_k{}^j = g^{ij} \quad ,$$

showing that there is a relation between the g^{ij}'s and the $\boldsymbol{\varepsilon}^i$'s analogous to that between the g_{ij}'s and the $\boldsymbol{\varepsilon}_i$'s,

$$\boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}_j = g_{ij} \qquad \text{and} \qquad \boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}^j = g^{ij} \quad .$$

From this, it follows that the g^{ij}'s are symmetric (since $\boldsymbol{\varepsilon}^i \cdot \boldsymbol{\varepsilon}^j = \boldsymbol{\varepsilon}^j \cdot \boldsymbol{\varepsilon}^i$). While we are at it, let's observe that there is a relation between the components of the metric and the Kronecker delta by simply noting that

$$\sum_{k=1}^{N} g^{ik} g_{kj} = \left[\mathbf{G}^{-1}\mathbf{G}\right]_{ij} = [\mathbf{I}]_{ij} \qquad \text{and} \qquad \sum_{k=1}^{N} g_{ik} g^{kj} = \left[\mathbf{G}\mathbf{G}^{-1}\right]_{ij} = [\mathbf{I}]_{ij} \quad .$$

So,

$$\sum_{k=1}^{N} g^{ik} g_{kj} = \delta^i{}_j \qquad \text{and} \qquad \sum_{k=1}^{N} g_{ik} g^{kj} = \delta_i{}^j \quad . \tag{11.5}$$

Co/Contra-variant Conversion for Vector Fields

Any vector field $\mathbf{F}$ on our space can be written using either its covariant or contravariant components,

$$\mathbf{F} = \sum_{i=1}^{N} F_i \,\boldsymbol{\varepsilon}^i = \sum_{j=1}^{N} F^j \,\boldsymbol{\varepsilon}_j \quad .$$

Using the relations

$$\boldsymbol{\varepsilon}_i = \sum_{j=1}^{N} g_{ij}\,\boldsymbol{\varepsilon}^j \quad , \quad \boldsymbol{\varepsilon}^i = \sum_{j=1}^{N} g^{ij}\,\boldsymbol{\varepsilon}_j \quad \text{and} \quad \boldsymbol{\varepsilon}^k \cdot \boldsymbol{\varepsilon}_j = \delta^k{}_j$$

(which we've already discussed), you can easily derive the formulas for computing the F_i's from the F^j's, along with the formulas for computing the F^j's from the F_i's:

?►*Exercise 11.3:* *Let $(F_1, F_2, \ldots, F_N)$ and $(F^1, F^2, \ldots, F^N)$ be, respectively, the co- and contravariant components of some vector field $\mathbf{F}$ with respect to some coordinate system. Verify the following:*

a: *The covariant and the contravariant components $\mathbf{F}$ are related to $\mathbf{F}$ by*

$$F^k = \mathbf{F} \cdot \boldsymbol{\varepsilon}^k \quad \text{and} \quad F_k = \mathbf{F} \cdot \boldsymbol{\varepsilon}_k \quad .$$

b: *The covariant and the contravariant components of $\mathbf{F}$ are related to each other by*

$$F^j = \sum_{k=1}^{N} g^{jk} F_k \quad \text{and} \quad F_i = \sum_{k=1}^{N} g_{ik} F^k \quad .$$

(Hint: Start with $\mathbf{F} = \sum_{k=1}^{N} F_k \boldsymbol{\varepsilon}^k$ and replace $\boldsymbol{\varepsilon}^k$ with the appropriate formula of $\boldsymbol{\varepsilon}_j$'s from above.)

11.5 Converting Between Two Coordinate Systems

Suppose we have a second coordinate system with reciprocal bases fields

$$\left\{(x^{1'}, x^{2'}, \ldots, x^{N'})\right\} \quad , \quad \left\{\boldsymbol{\varepsilon}_1{}', \boldsymbol{\varepsilon}_2{}', \ldots, \boldsymbol{\varepsilon}_N{}'\right\} \quad \text{and} \quad \left\{\boldsymbol{\varepsilon}^{1'}, \boldsymbol{\varepsilon}^{2'}, \ldots, \boldsymbol{\varepsilon}^{N'}\right\}$$

(where $\boldsymbol{\varepsilon}_i{}' = {}^{\partial \mathbf{r}}/_{\partial x^{i'}}$). Applying the chain rule, we get

$$\boldsymbol{\varepsilon}_i{}' = \frac{\partial \mathbf{r}}{\partial x^{i'}} = \sum_{j=1}^{N} \frac{\partial x^j}{\partial x^{i'}} \frac{\partial \mathbf{r}}{\partial x^j} = \sum_{j=1}^{N} \frac{\partial x^j}{\partial x^{i'}} \boldsymbol{\varepsilon}_j \quad .$$

So the formula for finding each $\boldsymbol{\varepsilon}_i'$ from the $\boldsymbol{\varepsilon}_j$'s is

$$\boldsymbol{\varepsilon}_i{}' = \sum_{j=1}^{N} \frac{\partial x^j}{\partial x^{i'}} \boldsymbol{\varepsilon}_j \quad \text{for} \quad i = 1, 2, \ldots, N \quad . \tag{11.6a}$$

Likewise

$$\boldsymbol{\varepsilon}_i = \sum_{j=1}^{N} \frac{\partial x^{j'}}{\partial x^i} \boldsymbol{\varepsilon}_j{}' \qquad \text{for} \quad i = 1, 2, \ldots, N \quad . \tag{11.6b}$$

Finding the relation between the $\boldsymbol{\varepsilon}^{i'}$'s and the $\boldsymbol{\varepsilon}^i$'s is a bit more tricky. Let

$$\boldsymbol{\varepsilon}^{i'} = \sum_{k=1}^{N} \alpha^i{}_k \boldsymbol{\varepsilon}^k$$

where the $\alpha^i{}_k$'s are to be determined. To be precise, they must be the (unique) scalar fields such that

$$\delta^i{}_j = \boldsymbol{\varepsilon}^{i'} \cdot \boldsymbol{\varepsilon}_j{}' = \left(\sum_{n=1}^{N} \alpha^i{}_n \boldsymbol{\varepsilon}^n \right) \cdot \left(\sum_{k=1}^{N} \frac{\partial x^k}{\partial x^{j'}} \boldsymbol{\varepsilon}_k \right) = \cdots = \sum_{k=1}^{N} \alpha^i{}_k \frac{\partial x^k}{\partial x^{j'}} \quad . \tag{11.7}$$

Now, observe that, because the $x^{k'}$'s are independent coordinates,

$$\frac{\partial x^{i'}}{\partial x^{j'}} = \text{rate } x^{i'} \text{ varies as } x^{j'} \text{ varies} = \begin{Bmatrix} 0 & \text{if} & j \neq i \\ 1 & \text{if} & i = j \end{Bmatrix} = \delta^i{}_j \quad .$$

This and the chain rule give us

$$\delta^i{}_j = \frac{\partial x^{i'}}{\partial x^{j'}} = \sum_{k=1}^{N} \frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^k}{\partial x^{j'}}$$

Comparing this with equation (11.7), we see that, for equation (11.7) to hold, we must have

$$\alpha^i{}_k = \frac{\partial x^{i'}}{\partial x^k} \quad .$$

Thus,

$$\boldsymbol{\varepsilon}^{i'} = \sum_{k=1}^{N} \frac{\partial x^{i'}}{\partial x^k} \boldsymbol{\varepsilon}^k \quad . \tag{11.8a}$$

Likewise

$$\boldsymbol{\varepsilon}^{i} = \sum_{k=1}^{N} \frac{\partial x^{i}}{\partial x^{k'}} \boldsymbol{\varepsilon}^{k'} \quad . \tag{11.8b}$$

With equations (11.6) and (11.8), you can now derive what may be the most important relations in tensor analysis.

?►*Exercise 11.4:* *Let* $\mathbf{F}$ *be a vector field.*

a: *Show that the contravariant components of* $\mathbf{F}$ *with respect to the* $\{(x^{1'}, x^{2'}, \ldots, x^{N'})\}$ *coordinate system are related to those with respect to the* $\{(x^1, x^2, \ldots, x^N)\}$ *coordinate system by*

$$F^{j'} = \sum_{i=1}^{N} \frac{\partial x^{j'}}{\partial x^i} F^i \quad . \tag{11.9}$$

(Hint: Start with $\mathbf{F} = \sum_{i=1}^{N} F^i \boldsymbol{\varepsilon}_i$ *and apply equation (11.6b).)*

b: *Show that the covariant components of* $\mathbf{F}$ *with respect to the* $\{(x^{1'}, x^{2'}, \ldots, x^{N'})\}$ *coordinate system are related to those with respect to the* $\{(x^1, x^2, \ldots, x^N)\}$ *coordinate system by*

$$F_i{}' = \sum_{j=1}^{N} \frac{\partial x^j}{\partial x^{i'}} F_j \quad . \tag{11.10}$$

Equation (11.9) is known as the "rank 1 contravariant transformation law", and equation (11.10) is known as the "rank 1 covariant transformation law". For some, these laws are the basic defining equations for tensors.

!▶*Example 11.3 (the gradient):* *Let* Ψ *be a scalar field with coordinate formulas* ψ *and* ψ' *with respect to the two coordinate systems. That is,*

$$\Psi(\boldsymbol{r}) = \psi\left(x^1, x^2, \ldots, x^N\right) \qquad \text{with} \quad \boldsymbol{r} \sim \left(x^1, x^2, \ldots, x^N\right)$$

and

$$\Psi(\boldsymbol{r}) = \psi'\left(x^{1'}, x^{2'}, \ldots, x^{N'}\right) \qquad \text{with} \quad \boldsymbol{r} \sim \left(x^{1'}, x^{2'}, \ldots, x^{N'}\right) \quad .$$

In example 11.2, we noted that the orthogonal coordinate system formula for the gradient of Ψ *could be written as*

$$\boldsymbol{\nabla}\Psi = \sum_{i=1}^{N} \frac{\partial \psi}{\partial x^i} \boldsymbol{\varepsilon}^i \quad ,$$

which meant that the covariant components of $\boldsymbol{\nabla}\Psi$ *are given by*

$$[\boldsymbol{\nabla}\Psi]_j = \frac{\partial \psi}{\partial x^j} \quad \text{when } \{(x^1, x^2, \ldots, x^N)\} \text{ is an orthogonal system.}$$

We suspect the same formula holds even if the system is not orthogonal.

To verify this suspicion, let $\{(x^1, \ldots, x^N)\}$ *be any orthogonal system (so the above formula for* $[\boldsymbol{\nabla}\Psi]_j$ *holds), and let* $\{(x^{1'}, \ldots, x^{N'})\}$ *be any other coordinate system, orthogonal or not. Applying the rank 1 covariant transformation law (i.e., equation (11.9)) and the classical chain rule, we get*

$$\begin{aligned}
[\boldsymbol{\nabla}\Psi]_i{}' &= \sum_{j=1}^{N} \frac{\partial x^j}{\partial x^{i'}} [\boldsymbol{\nabla}\Psi]_j = \sum_{j=1}^{N} \frac{\partial x^j}{\partial x^{i'}} \frac{\partial \psi}{\partial x^j} \\
&= \frac{\partial}{\partial x^{i'}} \left[\psi(x^1, x^2, \ldots, x^N)\right] \\
&= \frac{\partial}{\partial x^{i'}} \left[\psi'(x^{1'}, x^{2'}, \ldots, x^{N'})\right] = \frac{\partial \psi'}{\partial x^{i'}} \quad ,
\end{aligned}$$

confirming that

$$[\boldsymbol{\nabla}\Psi]_i = \frac{\partial \psi'}{\partial x^{i'}} \quad \text{using any coordinate system } \{(x^{1'}, x^{2'}, \ldots, x^{N'})\} \quad ,$$

and, thus, also confirming the suspicion expressed in exercise 11.2 that

$$\boldsymbol{\nabla}\Psi = \sum_{i=1}^{N} \frac{\partial \psi}{\partial x^i} \boldsymbol{\varepsilon}^i$$

*is a coordinate-*in*dependent formula for the gradient.*

11.6 So What Is A Tensor, Anyway?

I will give you two definitions: a good one, and the traditional one.

A Good (but Long) Definition of Tensors

Suppose we have an N-dimensional vector space $\mathcal{V}$. A *tensor* $\mathbf{T}$ is any single linear algebraic object — a scalar, a vector, a linear transformation of vectors, a linear transformation of linear transformations, etc. — defined on $\mathcal{V}$. The *rank* of $\mathbf{T}$ refers to the number of components of $\mathbf{T}$ with respect to any basis for $\mathcal{V}$. In particular:

$$\begin{aligned}
\mathbf{T} \text{ is rank } 0 \quad &\Longleftrightarrow \quad \mathbf{T} \text{ has } 1 = N^0 \text{ components (i.e., } \mathbf{T} \text{ is a } \textit{scalar}) \\
\mathbf{T} \text{ is rank } 1 \quad &\Longleftrightarrow \quad \mathbf{T} \text{ has } N^1 \text{ components (i.e., } \mathbf{T} \text{ is a } \textit{vector}) \\
\mathbf{T} \text{ is rank } 2 \quad &\Longleftrightarrow \quad \mathbf{T} \text{ has } N^2 \text{ components (e.g., } \mathbf{T} \text{ is a } \textit{linear transformation}) \\
&\ \vdots \\
\mathbf{T} \text{ is rank } m \quad &\Longleftrightarrow \quad \mathbf{T} \text{ has } N^m \text{ components}
\end{aligned}$$

Now suppose we have an N-dimensional space of positions. A *rank m tensor field* $\mathbf{T}$ (usually just called a "tensor") is just a rank m tensor-valued function of position. That is,

$$\mathbf{T}(\boldsymbol{p}) \;=\; \text{a rank } k \text{ tensor for the tangent vector space at } \boldsymbol{p} \quad .$$

So,

$$\begin{aligned}
\mathbf{T} \text{ is rank } 0 \text{ tensor field} \quad &\Longleftrightarrow \quad \mathbf{T}(\boldsymbol{p}) \text{ is a scalar for each position } \boldsymbol{p} \\
&\Longleftrightarrow \quad \mathbf{T} \text{ is a scalar field.} \\
\mathbf{T} \text{ is rank } 1 \text{ tensor field} \quad &\Longleftrightarrow \quad \mathbf{T}(\boldsymbol{p}) \text{ is a vector for each position } \boldsymbol{p} \\
&\Longleftrightarrow \quad \mathbf{T} \text{ is a vector field.} \\
&\ \vdots
\end{aligned}$$

Given a coordinate system and associated reciprocal basis fields

$$\left\{(x^1, x^2, \ldots, x^N)\right\} \quad , \quad \{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \quad \text{and} \quad \left\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\right\}$$

(where $\boldsymbol{\varepsilon}^k = {}^{\partial \boldsymbol{r}}/_{\partial x^k}$), the *covariant components* of $\mathbf{T}$ — denoted by T_i or T_{ij} or T_{ijk} or …, depending on the rank of $\mathbf{T}$ — are the components of $\mathbf{T}$ with respect to $\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\}$, while the *contravariant components* of $\mathbf{T}$ — denoted by T^i or T^{ij} or T^{ijk} or …, depending on the rank of $\mathbf{T}$ — are the components of $\mathbf{T}$ with respect to $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\}$.

This basically describes "what" tensors are. They are linear algebraic things defined on the tangent vector spaces. It can be shown (much as we have done for vector fields) that, if we have a second coordinate system with associated reciprocal basis fields

$$\left\{(x^{1'}, x^{2'}, \ldots, x^{N'})\right\} \quad , \quad \left\{\boldsymbol{\varepsilon}_{1'}, \boldsymbol{\varepsilon}_{2'}, \ldots, \boldsymbol{\varepsilon}_{N'}\right\} \quad \text{and} \quad \left\{\boldsymbol{\varepsilon}^{1'}, \boldsymbol{\varepsilon}^{2'}, \ldots, \boldsymbol{\varepsilon}^{N'}\right\}$$

(where $\boldsymbol{\varepsilon}^{k'} = {}^{\partial \mathbf{r}}/_{\partial x^{k'}}$), then the covariant components of a rank k tensor field $\mathbf{T}$ will satisfy the corresponding rank k transformation covariant law:

$$T_i{}' = \sum_{m=1}^{N} \frac{\partial x^m}{\partial x^{i'}} T_m \qquad \text{(rank 1)}$$

$$T_{ij}{}' = \sum_{m=1}^{N} \sum_{n=1}^{N} \frac{\partial x^m}{\partial x^{i'}} \frac{\partial x^n}{\partial x^{j'}} T_{mn} \qquad \text{(rank 2)}$$

$$T_{ijk}{}' = \sum_{m=1}^{N} \sum_{n=1}^{N} \sum_{o=1}^{N} \frac{\partial x^m}{\partial x^{i'}} \frac{\partial x^n}{\partial x^{j'}} \frac{\partial x^o}{\partial x^{k'}} T_{mno} \qquad \text{(rank 3)}$$

$$\vdots \quad ,$$

while its contravariant components will satisfy the corresponding rank k contravariant law of transformation:

$$T^{i'} = \sum_{m=1}^{N} \frac{\partial x^{i'}}{\partial x^m} T^m \qquad \text{(rank 1)}$$

$$T^{ij'} = \sum_{m=1}^{N} \sum_{n=1}^{N} \frac{\partial x^{i'}}{\partial x^m} \frac{\partial x^{j'}}{\partial x^n} T^{mn} \qquad \text{(rank 2)}$$

$$T^{ijk'} = \sum_{m=1}^{N} \sum_{n=1}^{N} \sum_{o=1}^{N} \frac{\partial x^{i'}}{\partial x^m} \frac{\partial x^{j'}}{\partial x^n} \frac{\partial x^{k'}}{\partial x^o} T^{mno} \qquad \text{(rank 3)}$$

$$\vdots \quad .$$

In addition, one can define and deal with mixed co- and contravariant components and the corresponding transformation laws.

Traditional Definition of Tensors

A rank 0 tensor (field) is a scalar field. For any positive integer m, a rank m $\left\{\begin{array}{c}\text{co-}\\ \text{contra-}\end{array}\right\}$ covariant tensor (field) consists of an infinite collection of sets of N^m scalar fields $\left\{\begin{array}{l}T_i \text{ or } T_{ij} \text{ or } \ldots\\ T^i \text{ or } T^{ij} \text{ or } \ldots\end{array}\right\}$ (with each set corresponding to a different coordinate system) that satisfy the rank m $\left\{\begin{array}{c}\text{co-}\\ \text{contra-}\end{array}\right\}$ variant transformation laws.

11.7 And What Is This Mysterious 'Metric' That Keeps Popping Up?

Simply put, the metric is our favorite "bilinear form", the dot product. To see this, you must first be told that a *bilinear form* $\mathcal{A}$ on a vector space $\mathcal{V}$ is a function which maps pairs of vectors into $\mathbb{R}$, and which is linear in each variable. That is, for every two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathcal{V}$, $\mathcal{A}(\mathbf{v},\mathbf{w})$ is a real number, and for every three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$, and every pair of real numbers α and β, we have

$$\mathcal{A}(\alpha\mathbf{v}+\beta\mathbf{w},\mathbf{u}) \;=\; \alpha\mathcal{A}(\mathbf{v},\mathbf{u}) \;+\; \beta\mathcal{A}(\mathbf{w},\mathbf{u})$$

and

$$\mathcal{A}(\mathbf{u},\alpha\mathbf{v}+\beta\mathbf{w}) \;=\; \alpha\mathcal{A}(\mathbf{u},\mathbf{v}) \;+\; \beta\mathcal{A}(\mathbf{u},\mathbf{w}) \quad .$$

Given any two bases $\mathcal{B}_1$ and $\mathcal{B}_2$ for the vector space, it can be shown that there is a matrix

$$\mathbf{A} \;=\; \mathbf{A}_{\mathcal{B}_2,\mathcal{B}_1}$$

such that

$$\mathcal{A}(\mathbf{v},\mathbf{w}) \;=\; \langle\mathbf{v}|_{\mathcal{B}_2}\;\mathbf{A}\;|\mathbf{w}\rangle_{\mathcal{B}_1} \quad .$$

The components of this matrix are called the *components of* $\mathcal{A}$ *with respect to bases* $\mathcal{B}_1$ *and* $\mathcal{B}_2$.

While we didn't discuss bilinear forms explicitly, there was one we used extensively. That was the *dot product* of vectors. Keep this in mind.

Now remember also, that we originally defined the covariant components of the metric, the g_{ij}'s, to satisfy

$$\left(\frac{ds}{dt}\right)^2 \;=\; \frac{d\boldsymbol{r}}{dt}\cdot\frac{d\boldsymbol{r}}{dt} \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} g_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{dx^i}{dt}g_{ij}\frac{dx^j}{dt} \quad .$$

Letting $\mathbf{v} = \dfrac{d\boldsymbol{r}}{dt}$ this becomes

$$\mathbf{v}\cdot\mathbf{v} \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} g_{ij}v^iv^j \;=\; \begin{bmatrix} v^1 & v^2 & \cdots & v^N\end{bmatrix}\mathbf{G}\begin{bmatrix} v^1 \\ v^2 \\ \vdots \\ v^N\end{bmatrix} \qquad \text{where} \quad [\mathbf{G}]_{ij} = g_{ij} \quad .$$

Further letting

$$\mathcal{B}_{\text{COV}} \;=\; \{\,\boldsymbol{\varepsilon}_1,\,\boldsymbol{\varepsilon}_2,\,\ldots,\,\boldsymbol{\varepsilon}_N\,\} \qquad \text{and} \qquad \mathcal{B}_{\text{CON}} \;=\; \{\,\boldsymbol{\varepsilon}^1,\,\boldsymbol{\varepsilon}^2,\,\ldots,\,\boldsymbol{\varepsilon}^N\,\}$$

this can be written as[4]

$$\mathbf{v}\cdot\mathbf{v} \;=\; \langle\mathbf{v}|_{\mathcal{B}_{\text{CON}}}\,\mathbf{G}\,|\mathbf{v}\rangle_{\mathcal{B}_{\text{CON}}} \quad .$$

[4] WARNING: In this section

$$\langle\mathbf{v}|_{\mathcal{B}} \;=\; |\mathbf{v}\rangle_{\mathcal{B}}{}^{\mathsf{T}} \quad .$$

This does not quite agree with our convention in previous chapters under which $\langle\mathbf{v}|_{\mathcal{B}}$ is the row matrix of the components of $\mathbf{v}$ with respect to the reciprocal basis.

This can be expanded using material developed in the previous several sections. Given two vector fields $\mathbf{v}$ and $\mathbf{w}$, we have

$$\mathbf{v}\cdot\mathbf{w} \;=\; \sum_{i=1}^{N} v^i w_i \;=\; \sum_{i=1}^{N} v^i \sum_{j=1}^{N} g_{ij} w^j \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} v^i g_{ij} w^j \;=\; \langle \mathbf{v}|_{\mathcal{B}_{\text{CON}}} \mathbf{G}\, |\mathbf{w}\rangle_{\mathcal{B}_{\text{CON}}}$$

Likewise, you can verify that

$$\mathbf{v}\cdot\mathbf{w} \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} v_i g^{ij} w_j \;=\; \langle \mathbf{v}|_{\mathcal{B}_{\text{CON}}} \mathbf{G}^{-1}\, |\mathbf{w}\rangle_{\mathcal{B}_{\text{CON}}}$$

Now you can pretty well see what "the metric" really is — it is the bilinear form $\mathcal{G}$ which is simply the dot product,

$$\mathcal{G}(\mathbf{v},\mathbf{w}) \;=\; \mathbf{v}\cdot\mathbf{w} \quad .$$

The covariant components of the metric are just the components of this bilinear form with respect to $\mathcal{B}_{\text{CON}}$ and the contravariant components are just the components with respect to $\mathcal{B}_{\text{COV}}$. That is,

$$\mathbf{G} \;=\; \mathbf{G}_{\mathcal{B}_{\text{CON}},\mathcal{B}_{\text{CON}}} \qquad \text{and} \qquad \mathbf{G}^{-1} \;=\; \mathbf{G}_{\mathcal{B}_{\text{COV}},\mathcal{B}_{\text{COV}}} \quad .$$

Moreover, if you think about it, we also have

$$\mathbf{G}_{\mathcal{B}_{\text{CON}},\mathcal{B}_{\text{COV}}} \;=\; \mathbf{I} \;=\; \mathbf{G}_{\mathcal{B}_{\text{COV}},\mathcal{B}_{\text{CON}}}$$

where, as you should recall, $\mathbf{I}$ is the $N\times N$ identity matrix.

11.8 The Metric and the Christoffel Symbols

Let us assume we have an N-dimensional space with coordinate system

$$\left\{x^1, x^2, \ldots, x^N\right\} \quad .$$

At each point, we also have the corresponding reciprocal pair of bases

$$\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_N\} \qquad \text{and} \qquad \left\{\boldsymbol{\varepsilon}^1, \boldsymbol{\varepsilon}^2, \ldots, \boldsymbol{\varepsilon}^N\right\}$$

where $\boldsymbol{\varepsilon}_j \;=\; {}^{\partial \mathbf{x}}/_{\partial x^j}$. We also have the co- and contravariant components of the metric

$$g_{ij} \;=\; \boldsymbol{\varepsilon}_i\cdot\boldsymbol{\varepsilon}_j \qquad \text{and} \qquad g^{ij} \;=\; \boldsymbol{\varepsilon}^i\cdot\boldsymbol{\varepsilon}^j \quad ,$$

and we know that

$$\frac{\partial \boldsymbol{\varepsilon}_i}{\partial x^j} \;=\; \sum_{n=1}^{N} \Gamma_{ij}^n \boldsymbol{\varepsilon}_n \quad .$$

Thus,

$$\begin{aligned}\frac{\partial g_{ij}}{\partial x^k} \;&=\; \frac{\partial}{\partial x^k}\left[\boldsymbol{\varepsilon}_i\cdot\boldsymbol{\varepsilon}_j\right]\\ &=\; \frac{\partial \boldsymbol{\varepsilon}_i}{\partial x^k}\cdot\boldsymbol{\varepsilon}_j \;+\; \boldsymbol{\varepsilon}_i\cdot\frac{\partial \boldsymbol{\varepsilon}_j}{\partial x^k}\end{aligned}$$

$$= \left[\sum_{n=1}^{N} \Gamma_{ik}^{n} \boldsymbol{\varepsilon}_n\right] \cdot \boldsymbol{\varepsilon}_j \;+\; \boldsymbol{\varepsilon}_i \cdot \left[\sum_{n=1}^{N} \Gamma_{jk}^{n} \boldsymbol{\varepsilon}_n\right]$$

$$= \sum_{n=1}^{N} \Gamma_{ik}^{n} \boldsymbol{\varepsilon}_n \cdot \boldsymbol{\varepsilon}_j \;+\; \sum_{n=1}^{N} \Gamma_{jk}^{n} \boldsymbol{\varepsilon}_i \cdot \boldsymbol{\varepsilon}_n \quad .$$

Rewriting the dot products as the covariant components of the metric, this becomes

$$\frac{\partial g_{ij}}{\partial x^k} = \sum_{n=1}^{N} \Gamma_{ik}^{n} g_{nj} + \sum_{n=1}^{N} \Gamma_{jk}^{n} g_{ni} \qquad \text{for} \quad i = 1, \ldots N \,,\; j = 1, \ldots N \,,\; k = 1, \ldots N \quad . \tag{11.11}$$

This is a linear system of N^3 linear equations for computing the N^3 partial derivatives of the components of the metric from the components of the metric and the N^3 Christoffel symbols. But if you already have coordinate formulas for the covariant components of the metric, you could just compute the above partial derivatives from those formulas without using the above system. You would not even need to know the formulas for the Christoffel symbols.

On the other hand, we can view equation set (11.11) as a linear system with the Christoffel symbols as the unknowns, and, with luck, we can "invert" this system to get a set of linear formulas for the N^3 Christoffel symbols in terms of the components of the metric and their partial derivatives. In theory, this inversion could be done by converting system (11.11) to a giant "matrix/vector" equation (after eliminating the redundancies due to symmetries in the elements), and then finding the inverse of that matrix. But that is far too difficult. Instead we will use a more elementary approach which clearly demonstrates the devious cleverness of the mathematicians who first derived this.

To simplify our computation, we first rewrite each equation in system (11.11) as

$$\frac{\partial g_{ij}}{\partial x^k} = A(i, k; j) + A(j, k; i) \tag{11.12a}$$

where

$$A(\alpha, \beta; \gamma) = \sum_{n=1}^{N} \Gamma_{\alpha\beta}^{n} g_{n\gamma} \quad . \tag{11.12b}$$

(The $A(\alpha, \beta; \gamma)$'s are called the Christoffel symbols of the first kind, and are more traditionally denoted by $[\alpha\beta, \gamma]$. What we have been calling the Christoffel symbols are actually the Christoffel symbols of the second kind.)

Observe that, because of a symmetry in the Christoffel symbols of second kind , there is a corresponding symmetry in the A's :

$$A(\beta, \alpha; \gamma) = \sum_{n=1}^{N} \Gamma_{\beta\alpha}^{n} g_{n\gamma} = \sum_{n=1}^{N} \Gamma_{\alpha\beta}^{n} g_{n\gamma} = A(\alpha, \beta; \gamma) \quad .$$

Also, because of the relation between the co- and contravariant components of the metric tensor,

$$\sum_{\gamma=1}^{N} g^{\mu\gamma} A(\alpha, \beta; \gamma) = \sum_{\gamma=1}^{N} g^{\mu\gamma} \sum_{n=1}^{N} \Gamma_{\alpha\beta}^{n} g_{n\gamma}$$

$$= \sum_{n=1}^{N} \Gamma_{\alpha\beta}^{n} \sum_{\gamma=1}^{N} g^{\mu\gamma} g_{\gamma n} = \sum_{n=1}^{N} \Gamma_{\alpha\beta}^{n} \delta^{\mu}{}_{n} \quad .$$

So

$$\Gamma^{\mu}_{\alpha\beta} \;=\; \sum_{\gamma=1}^{N} g^{\mu\gamma} A(\alpha, \beta; \gamma) \quad . \tag{11.13}$$

Now observe that, repeatedly using equation (11.12a) and symmetries, we get

$$\begin{aligned}
A(\alpha, \beta; \gamma) &= \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} - A(\gamma, \beta; \alpha) \\
&= \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} - A(\beta, \gamma; \alpha) \\
&= \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} - \left[\frac{\partial g_{\beta\alpha}}{\partial x^{\gamma}} - A(\alpha, \gamma; \beta) \right] \\
&= \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\gamma}} + A(\gamma, \alpha; \beta) \\
&= \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\gamma}} + \left[\frac{\partial g_{\gamma\beta}}{\partial x^{\alpha}} - A(\beta, \alpha; \gamma) \right] \\
&= \frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\gamma}} + \frac{\partial g_{\gamma\beta}}{\partial x^{\alpha}} - A(\beta, \alpha; \gamma) \quad .
\end{aligned}$$

Cutting out the middle and solving the above for $A(\alpha, \beta; \gamma)$ gives us

$$A(\alpha, \beta; \gamma) \;=\; \frac{1}{2} \left[\frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} + \frac{\partial g_{\beta\gamma}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\gamma}} \right] \quad .$$

Combining this with formula (11.13) for $\Gamma^{\mu}_{\alpha\beta}$ then gives us the main result of this section:

Theorem 11.2
Given any coordinate system $\{x^1, x^2, \ldots, x^N\}$ for any N-dimensional space, the corresponding Christoffel symbols of the second kind can be computed from the co- and contravariant components of the metric via

$$\Gamma^{\mu}_{\alpha\beta} \;=\; \frac{1}{2} \sum_{\gamma=1}^{N} g^{\mu\gamma} \left[\frac{\partial g_{\alpha\gamma}}{\partial x^{\beta}} + \frac{\partial g_{\beta\gamma}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\gamma}} \right] \tag{11.14}$$
$$\text{for} \quad \alpha = 1, \ldots N \,,\; \beta = 1, \ldots N \,,\; \mu = 1, \ldots N \quad .$$