

Tensor Calculus

(à la Speedy Gonzales)

The following is a lightning introduction to Tensor Calculus. The presentation in *Mathematical Methods for Physicists* by G. Arfken is misleading. It does not distinguish between co- and contra-variant (cotangent and tangent) vectors in $\frac{7}{9}$ of Chapter 3. Sections 3.8 and 3.9 finally do introduce “Noncartesian Tensors”. This is about as pedagogical as (1) dividing the entire fauna into Hippopotami and “Nonhippopotamous animals”, then (2) spending more than $\frac{3}{4}$ of your time studying these water beasts and thereupon (3) less than $\frac{1}{4}$ of your time in generalizing the graceful characteristics of this species to the remainder of the animal kingdom.

1. Why and What With All This Jazz!?!?

First of all, the whole purpose of the Tensor Calculus is to provide a unified framework for dealing with quantities (and their derivatives and integrals) which depend on several variables *and* need several numbers (a.k.a. *components*) per point in space to be fully specified. Whilst the number of these variables (generalized coordinates) and components is less than four, the usual human imagination and geometric intuition, and a little pedantry is often quite satisfactory. However, precious few are the problems which involve only this small number of variables¹⁾.

The classic *Tensor Calculus* by J.L. Synge and A. Schild (Dover, New York, 1978) is wholeheartedly recommended, because it is self-contained, inexpensive (about \$7.–) and because it covers some easy application to Classical Mechanics, Hydrodynamics, Elasticity, Electromagnetism, Relativity, . . . Finally, it treats Cartesian coordinates as a (very) special case, which they indeed are.



Having learned (and having been intimidated by) some “vector calculus”, with all the many “vector identities”, (multiple) curls, divergences and gradients, contour,- surface- and volume-integrals. . . you may rightly be asking why tensors? (Never mind what they are. . .)

Perhaps the most honest answer is “because they exist”. Examples are not as abundant in daily experience as vectors and scalars, but they exist: the *metric*, which is used to define the *line element* (the differential along a given but arbitrary curve); the electromagnetic *stress-tensor*; the conductivity in an anisotropic medium; the generalized Young modulus which correlates shears and stresses with the forces which created them; . . .

While it is admittedly comfortable to start with the (hopefully) well understood and already familiar Cartesian coordinate system, many things become literally trivial and

¹⁾ Even a single billiard (pool) ball requires five or occasionally six coordinates: three for rotations (spin) and two (three if it jumps) for translation! As for the whole game. . .

obvious *only in sufficient generality*²⁾. This will be one major motivation for the study of tensors in general.

The other motivation is more in the *method* than in the madness of it; that is, to some extent, the focus on the particular technique and squeezing as much purchase out of it as possible will be an end unto itself. The technique relies on the fact that whatever coordinates are being used are merely a figment of our mathematical *description*, and not really essential; therefore, physically meaningful statements (equations, laws, ...) should be *independent of any choice of coordinates* whatsoever. Now, all physically meaningful statements (equations, laws, ...) are made in terms of physically meaningful quantities (temperature, force, acceleration, coefficient of viscosity, electrostatic field, ...). Thus, we must first learn how these “building blocks” change when we (arbitrarily) change the coordinate system, so as to be able to put these together into meaningful statements.

Thereupon, writing down physical laws will still not be as easy as building with A-B-C blocks, but ill-formed candidates will crumble most obviously. In fact, several of the “integration order reducing formulae”, such as Gauss’s and Stoke’s laws can be proven by simply observing that no other expression can be constructed with the required number of derivatives and integrations, and the integrand at hand, subject to the fact that the left-hand-side and the right-hand-side must be quantities of the same type (scalar, vector, ...). Many of the arguments of tensor calculus come as an evolution of this principle.

1.1. Coordinate systems are free to choose!

The fundamental idea in tensor calculus is the transformation of variables. Given a collection of linearly independent (generalized) coordinates³⁾ $x^i, i = 1, 2, \dots, n$ to *span* (describe, parametrize, coordinatize) an n -dimensional space, any point in this n -dimensional space is unambiguously given by as the ordered n -tuple (x^1, x^2, \dots, x^n) .

No one may stop us from adopting another collection of (perhaps more suitable or simply nicer) coordinates

$$\tilde{x}^i = \tilde{x}^i(x^1, x^2, \dots, x^n), \quad i = 1, 2, \dots, n. \quad (1.1)$$

Clearly, every point in the same n -dimensional space now must be unambiguously representable by an ordered n -tuple $(\tilde{x}^1, \tilde{x}^2, \dots, \tilde{x}^n)$. To communicate back and forth with all those who prefer the original coordinate system $\{x^i\}$ and ensure that the two coordinate systems are both worth the paper on which they are written, the n equations (1.1) must be *invertible*. As everyone should know by now, that means that the *Jacobian* of the transformation

$$\mathcal{J}\left(\frac{\partial(\tilde{x}^1, \dots, \tilde{x}^n)}{\partial(x^1, \dots, x^n)}\right) \stackrel{\text{def}}{=} \det \begin{pmatrix} \frac{\partial \tilde{x}^1}{\partial x^1} & \dots & \frac{\partial \tilde{x}^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{x}^n}{\partial x^1} & \dots & \frac{\partial \tilde{x}^n}{\partial x^n} \end{pmatrix} \quad (1.2)$$

is non-zero. We assume this from now on.

²⁾ This is very, very much like the roundness of the Earth, which is difficult to comprehend unless one is willing to take a global. But, when one does, things like the Coriolis “force” and its geographical effects become obvious.

³⁾ Please, note that x^i is not the i^{th} power of the variable x , but the i^{th} variable in the collection. As both *sub*- and *superscripts* do occur eventually, there is no reason to obstinately insist on *subscripting* the coordinate variables.

Note the matrix in Eq. (1.2), called imaginatively the ‘transformation matrix’. The name is indeed descriptive, as it is used—you guessed it—to transform things from one set of coordinates to another. Consider the differential of each of the n equations (1.1):

$$d\tilde{x}^i = \sum_{j=1}^n \left(\frac{\partial \tilde{x}^i}{\partial x^j} \right) dx^j, \quad \text{for each } i = 1, 2, \dots, n. \quad (1.3)$$

Summations such as this occur all the time and present a tedious nuisance. Note also that the summation index, here j , occurs once as a superscript (on dx^j) and once as a subscript on $\frac{\partial}{\partial x^j}$. Of course, a superscript on something in the denominator counts as a subscript; also, a subscript on something in the denominator counts as a superscript. We therefore adopt the

Einstein Summation Convention: When an index is repeated—once as a superscript and once as a subscript—a summation over this index is understood, the range of the summation being $1, 2, \dots, n$.

Thus, we can rewrite Eq. (1.3) as $d\tilde{x}^i = \frac{\partial \tilde{x}^i}{\partial x^j} dx^j$.

The coordinates x^i , $i = 1, 2, \dots, n$ are components of the radius vector $\vec{r} = x^i \hat{e}_i$ (remember summation?), the one which points from the origin $(0, 0, \dots, 0)$ to the point (x^1, x^2, \dots, x^n) . Notice that while the coordinates x^i (components of the vector \vec{r}) have a *superscript*, the unit vectors \hat{e}_i have a *subscript*. What is that all about?

Well, note that the components of a gradient have *subscripts*:

$$[\vec{\nabla} f]_i = \frac{\partial f}{\partial x^i} \quad (1.4)$$

The textbook by Arfken is misleading by not distinguishing a vector with *contra-variant* components, such as x^i , from *co-variant ones*, such as $\frac{\partial f}{\partial x^i}$. Let’s see how these differ.

Let \tilde{x}^i and x^i be coordinates in two different coordinate systems. Then, for example, (note the summation)

$$d\tilde{x}^i = \left(\frac{\partial \tilde{x}^i}{\partial x^j} \right) dx^j, \quad \text{foreach } i = 1, 2, \dots, n, \quad (1.5)$$

using the chain-rule. Straightforward, right? Well, equally straightforward should be the case of the gradient of a scalar function $f(x^i) = \tilde{f}(\tilde{x}^i)$, for which (note the summation)

$$\left(\frac{\partial \tilde{f}}{\partial \tilde{x}^i} \right) = \left(\frac{\partial x^j}{\partial \tilde{x}^i} \right) \left(\frac{\partial f}{\partial x^j} \right) \quad \text{for each } i = 1, 2, \dots, n. \quad (1.6)$$

Note the *absolutely crucial fact*, that the components of $d\vec{r}$ transform with $\left(\frac{\partial \tilde{x}^i}{\partial x^j} \right)$, while the components of $\vec{\nabla} f$ transform with $\left(\frac{\partial x^j}{\partial \tilde{x}^i} \right)$!!!

These transformation factors look quite opposite to each other. In fact, when viewed as matrices, they are inverses of one another. Writing all n equations (1.5), we have

$$\begin{bmatrix} d\tilde{x}^1 \\ \vdots \\ d\tilde{x}^n \end{bmatrix} = \begin{bmatrix} \frac{\partial \tilde{x}^1}{\partial x^1} & \cdots & \frac{\partial \tilde{x}^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{x}^n}{\partial x^1} & \cdots & \frac{\partial \tilde{x}^n}{\partial x^n} \end{bmatrix} \begin{bmatrix} dx^1 \\ \vdots \\ dx^n \end{bmatrix}, \quad (1.7)$$

while

$$\begin{bmatrix} \frac{\partial \tilde{f}}{\partial \tilde{x}^1} \\ \vdots \\ \frac{\partial \tilde{f}}{\partial \tilde{x}^n} \end{bmatrix} = \begin{bmatrix} \frac{\partial x^1}{\partial \tilde{x}^1} & \cdots & \frac{\partial x^1}{\partial \tilde{x}^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial \tilde{x}^1} & \cdots & \frac{\partial x^n}{\partial \tilde{x}^n} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x^1} \\ \vdots \\ \frac{\partial f}{\partial x^1} \end{bmatrix}. \quad (1.8)$$

In fact, leaving the index i to run freely over $1, 2, \dots, n$ in Eqs. (1.5) and (1.6), those equations are equivalent to the above matrix equations. Also, it should be obvious that the matrices $\left[\frac{\partial \tilde{x}^i}{\partial x^j}\right]$ and $\left[\frac{\partial x^j}{\partial \tilde{x}^i}\right]$ are inverse to each other (do the calculation!):

$$\begin{bmatrix} \frac{\partial \tilde{x}^1}{\partial x^1} & \cdots & \frac{\partial \tilde{x}^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \tilde{x}^n}{\partial x^1} & \cdots & \frac{\partial \tilde{x}^n}{\partial x^n} \end{bmatrix} \begin{bmatrix} \frac{\partial x^1}{\partial \tilde{x}^1} & \cdots & \frac{\partial x^1}{\partial \tilde{x}^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial \tilde{x}^1} & \cdots & \frac{\partial x^n}{\partial \tilde{x}^n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (1.9)$$

This same, rewritten “in the index notation”, becomes:

$$\frac{\partial \tilde{x}^i}{\partial x^j} \frac{\partial x^j}{\partial \tilde{x}^k} = \frac{\partial \tilde{x}^i}{\partial \tilde{x}^k} = \delta_k^i, \quad \delta_k^i = \begin{cases} 1 & i = k, \\ 0 & i \neq k, \end{cases} \quad (1.10)$$

since the n variables \tilde{x}^i are linearly independent. Of course, it is also true that $\frac{\partial x^i}{\partial \tilde{x}^j} \frac{\partial \tilde{x}^j}{\partial x^k} = \delta_k^i$. The quantity δ_k^i is called the Kronecker symbol. To save spacetime and paper, we’ll never write out things like (1.9) again.

—○—

Thus we conclude, not all vectors transform equally with respect to an arbitrary change of coordinates. Some will transform as in (1.6), and some as in (1.5). This turns out to exhaust all possibilities, and so we in general have (1.5)-like and (1.6)-like vectors. Historically, the following names became adopted:

Definition.

A vector is **contra-variant** if its components transform oppositely from the way those of $\vec{\nabla}$ do.

A vector is **co-variant** if its components transform the way those of $\vec{\nabla}$ do.

—○—

As frivolous as it may seem, this notation easily provides for the following.

A vector $\vec{A} = A^i \hat{e}_i$ (describing the wind, the rotation of the Earth around the Sun, the Gravitational field, electrostatic field, ...) couldn’t care less which coordinates we choose to describe it. Thus, it must be true that $\vec{\tilde{A}} = \vec{A}$. Indeed:

$$\begin{aligned} \vec{\tilde{A}} = \tilde{A}^i \hat{e}_i &= \left[\left(\frac{\partial \tilde{x}^i}{\partial x^j} \right) A^j \right] \left[\left(\frac{\partial x^k}{\partial \tilde{x}^i} \right) \hat{e}_k \right] \\ &= \left[\left(\frac{\partial \tilde{x}^i}{\partial x^j} \right) \left(\frac{\partial x^k}{\partial \tilde{x}^i} \right) \right] A^j \hat{e}_k = \delta_j^k \hat{e}_k A^j = A^k \hat{e}_k = \vec{A}, \end{aligned} \quad (1.11)$$

where we have used that (recall summation) $\delta_j^k A^j = A^k$ (why?), and Eq. (1.9).

Just the same, if we have a contravariant vector \vec{A} , (with components A^i) and another covariant vector \vec{B} (with components B_i), we can form a product ($A^i B_i$), which does not transform. Since precisely that—*no transformation*—is the key property of what is called a *scalar*, the product ($A^i B_i$) has earned its name—the *scalar product*.

More Names.

One also says that two indices are “contracted”: in $A^i \hat{e}_i$, the two copies of i are contracted and there is an implicit summation with i running over $1, 2, \dots, n$. A contracted index is also called a “dummy” index. It is sort-of the discrete version of a variable over which one has integrated.

1.2. More Products and More Indices

Suppose we are given two contravariant vectors, \vec{A} and \vec{B} , with components A^i, B^j . Consider the product $A^i B^j$, with both i and j left to run freely $i, j = 1, 2, \dots, n$. Clearly, there are n^2 such quantities: $A^1 B^1, A^1 B^2, A^1, B^3, \dots, A^n B^n$. How does such a composite quantity $C^{ij} = A^i B^j$ transform?

Straightforwardly:

$$\tilde{C}^{ij} = \tilde{A}^i \tilde{B}^j = \frac{\partial \tilde{x}^i}{\partial x^k} \frac{\partial \tilde{x}^j}{\partial x^l} A^k B^l = \frac{\partial \tilde{x}^i}{\partial x^k} \frac{\partial \tilde{x}^j}{\partial x^l} C^{kl} . \quad (1.12)$$

More generally, a quantity that has p superscripts and q subscripts may transform simply as a product of p contravariant vector components and q covariant vector components; that is, as

$$\tilde{T}_{j_1 \dots j_q}^{i_1 \dots i_p} = \frac{\partial \tilde{x}^{i_1}}{\partial x^{k_1}} \dots \frac{\partial \tilde{x}^{i_p}}{\partial x^{k_p}} \frac{\partial x^{l_1}}{\partial \tilde{x}^{j_1}} \dots \frac{\partial x^{l_q}}{\partial \tilde{x}^{j_q}} T_{l_1 \dots l_q}^{k_1 \dots k_p} . \quad (1.13)$$

Definition.

Any quantity like $T_{l_1 \dots l_q}^{k_1 \dots k_p}$, which transforms according to Eq. (1.13) for some p, q , is called a **tensor**.

The total number of transformation factors (=number of free indices) acting on a tensor $T_{l_1 \dots l_q}^{k_1 \dots k_p}$ is called the **rank** of the tensor (= $p+q$).

The **type** of a tensor is the pair of numbers (p, q) , specifying the transformations matrices $\frac{\partial \tilde{x}}{\partial x}$ and $\frac{\partial x}{\partial \tilde{x}}$ separately.

—○—

Section 3.3 introduces what is known as the “quotient rule” and is likely to confuse. Here’s (one possible version of) the corrected list:

$$K^i A_i = B \quad (3.29a')$$

$$K_j^i A_i = B_j \quad (3.29b')$$

$$K^{ij} A_{jk} = B_k^i \quad (3.29c')$$

$$K_{kl}^{ij} A_{ij} = B_{kl} \quad (3.29d')$$

$$K_{ij} A_k = B_{ijk} \quad (3.29e')$$

Note that the contracted indices do not appear on the other side, while all the free ones do. This is precisely the “quotient rule” (not that anything is being quotiented): the left- and the right-hand side of any equation must transform the same and this can easily be checked to be true for any of these or similar expressions. Sometimes, this is also called the “index conservation rule”, and is simply an extension of the old saw about being careful not to equate Apples with PC’s.

Let us check Eq. (3.29b’). In the twiddled coordinate system, it reads $\tilde{K}_j^i \tilde{A}_i = \tilde{B}_j$. Since \tilde{B}_j are components of a covariant vector, we have

$$\tilde{K}_j^i \tilde{A}_i = \tilde{B}_j = \frac{\partial x^l}{\partial \tilde{x}^j} B_l \stackrel{(3.29b')}{=} \frac{\partial x^l}{\partial \tilde{x}^j} (K_l^m A_m), \quad (1.14)$$

in using Eq. (3.29b’), the free index was l and we labeled the dummy index by m (it can be anything you please, as long as it cannot get confused with some other index⁴⁾).

Now, A_m is also covariant; transforming it *back* to the twiddled system and moving r.h.s. to the left produces

$$\tilde{K}_j^i \tilde{A}_i - \frac{\partial x^l}{\partial \tilde{x}^j} K_l^m \frac{\partial \tilde{x}^i}{\partial x^m} \tilde{A}_i = 0, \quad (1.15)$$

which is easily rewritten as

$$\left[\tilde{K}_j^i - \frac{\partial x^l}{\partial \tilde{x}^j} K_l^m \frac{\partial \tilde{x}^i}{\partial x^m} \right] \tilde{A}_i = 0. \quad (1.16)$$

As this must be true for any \tilde{A}_i (we only used its transformation properties, not the direction or magnitude), it must be that

$$\tilde{K}_j^i = \frac{\partial x^l}{\partial \tilde{x}^j} \frac{\partial \tilde{x}^i}{\partial x^m} K_l^m, \quad (1.17)$$

which says that K_j^i is a rank-2 tensor, of type (1,1), as we could have read off directly from K_j^i having two free indices, one superscript and one subscript—thus, a little care with the “index notation” makes the calculation (1.14)–(1.17) trivial.

Finally, note that by keeping consistently track of upper and lower indices (super- and sub-scripts) of a tensor, we know *precisely* how it transforms under an *arbitrary* change of coordinates. Also, the footnote on p.127 is unnecessary; you don’t need to lose your mind over deciding which “cosine a_{jl} ” to use.

—○—

Consider the i^{th} component of the gradient of a scalar, $[\vec{\nabla} f]_i = \frac{\partial f}{\partial x^i}$. Since both the i^{th} component of $\vec{\nabla} f$ and \hat{e}_i are components of a covariant vector, how can we contract them? Recall that this issue is not new: the components of $d\vec{r}$ are contravariant (dx^i), yet we *can* calculate its magnitude,

$$ds^2 \stackrel{\text{def}}{=} d\vec{r} \cdot d\vec{r} = g_{ij} dx^i dx^j, \quad (1.18)$$

⁴⁾ A dummy index is very much like the integration variable in a definite integral; do not ever use a symbol which was already present in the expression for a new dummy index!!!

where g_{ij} is called the *metric*. ds^2 is taken to mean the (physically measurable) distance between two (infinitesimally near) points and should be independent of our choice of coordinates; ds^2 is defined to be a scalar, hence the scalar product. For the left- and the right-hand side of this defining equation to make sense, g_{ij} must transform as a rank-2 covariant tensor. Moreover, $g_{ij} = g_{ji}$, since we can swap dx^i and dx^j and then relabeling $i \leftrightarrow j$ without ever changing the overall sign:

$$\begin{aligned} g_{ij} dx^i dx^j &= g_{ij} dx^j dx^i = g_{ji} dx^i dx^j , \\ [g_{ij} - g_{ji}] dx^i dx^j &= 0 , \end{aligned} \tag{1.19}$$

whence $g_{ij} = g_{ji}$.

g_{ij} being a matrix, we define g^{ij} to be the matrix-inverse of g_{ij} . That is, g^{ij} is defined to be that matrix which satisfies

$$g^{ij} g_{jk} = \delta_k^i , \quad \text{and} \quad g_{ij} g^{jk} = \delta_i^k . \tag{1.20}$$

The uniqueness of g^{ij} is an elementary fact of matrix algebra. The only minor point we need to make is that g_{ij} is also a function of space, $g_{ij} = g_{ij}(x^1, \dots, x^n)$, while δ_k^i is a constant. Clearly, therefore, $g^{ij} = g^{ij}(x^1, \dots, x^n)$ is then such a matrix-valued function that the relations (1.20) hold point-by-point in all of the (x^1, \dots, x^n) -space.

Having introduced g^{ij} , it is easy to write

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x^i} \right) g^{ij} \hat{e}_j = \left(\frac{\partial f}{\partial x^i} \right) \hat{e}^i , \tag{1.21}$$

with

$$\hat{e}^i \stackrel{\text{def}}{=} g^{ij} \hat{e}_j . \tag{1.22}$$

More generally, just like $d\vec{r} \cdot d\vec{r} = g_{ij} dx^i dx^j$, we have

$$\begin{aligned} \vec{A} \cdot \vec{B} &= g_{ij} A^i B^j , \\ \vec{C} \cdot \vec{D} &= g^{ij} C_i D_j , \\ \vec{A} \cdot \vec{C} &= A^i C_i , \end{aligned} \tag{1.23}$$

and so on, for any two contravariant vectors \vec{A}, \vec{B} and any two covariant vectors \vec{C}, \vec{D} . The combination $(g_{ij} A^i) = A_j$ transforms as a covariant tensor; the index on A^i has been *lowered*. Similarly, the combination $g^{ij} C_i = C^j$ transforms as a contravariant vector; the index on C_i has been *raised*. Thus, upon *rising* and *lowering* indices, contractions are performed just like in Eqs. (3.29')—as corrected above.

Of course, even if we have tensors of higher rank, the metric g_{ij} or the inverse-metric g^{ij} is used to contract indices. For example, the tensors T^{ij}, U^{klm} may be contracted using the metric once:

$$g_{ik} T^{ij} U^{klm} = [\mathbf{T} \cdot \mathbf{U}]^{jlm} , \tag{1.24}$$

or twice

$$g_{ik} g_{jl} T^{ij} U^{klm} = [\mathbf{T} \cdot \mathbf{U}]^m . \tag{1.25}$$

Clearly, at this point, the $\vec{\cdot}$ -and- \cdot notation becomes very confusing and is best abandoned. (For tensors of rank 5, you'd need to stack five arrows atop each other.)

1.3. The cross-product

The formulae (1.23) provide a perfectly general definition of the scalar product of any two vectors; the expressions in (1.23) are invariant under absolutely any change of coordinates, and in any number of dimensions.

How about the cross-product then? Well, recall the ‘primitive’ definition $\vec{A} \times \vec{B} = \hat{n}|\vec{A}||\vec{B}|\sin\theta$, where θ is the angle between \vec{A} and \vec{B} , \hat{n} is the unit vector perpendicular to both \vec{A} and \vec{B} , and chosen such that the triple $\vec{A}, \vec{B}, \hat{n}$ forms a right-hand triple. Already this reveals that *the cross-product exists only in three dimensions!* In two dimensions, there can be no \hat{n} (in two dimensions, there is no third direction!), and in n dimensions, there is are $n-2$ linearly independent \hat{n} ’s all of which orthogonal to the two given vectors.

Without much ado, the standard determinant formula for the cross-product of two three-dimensional covariant vectors is matched with the following index-notation formula:

$$(\vec{A} \times \vec{B})^i = \epsilon^{ijk} A_j B_k, \quad (1.26)$$

where

$$\epsilon^{ijk} \stackrel{\text{def}}{=} \begin{cases} 1 & i, j, k \text{ is an even permutation of } 1, 2, 3; \\ -1 & i, j, k \text{ is an odd permutation of } 1, 2, 3; \\ 0 & \text{otherwise,} \end{cases} \quad (1.27)$$

is the Levi-Civita symbol, also called the totally antisymmetric symbol, or the alternating symbol. Indeed, $\epsilon^{ijk} = -\epsilon^{ikj} = -\epsilon^{jik} = -\epsilon^{kji} = +\epsilon^{jki} = +\epsilon^{kij}$, for $i, j, k = 1, 2, 3$. Note that the cross product of two covariant vectors produced a contravariant one (1.26). Having learned however the trick that contraction with the metric (*i.e.*, the inverse-metric) lowers (*i.e.*, raises) indices, we can also define

$$(\vec{A} \times \vec{C})^i = \epsilon^{ijk} A_j g_{kl} C^l, \quad (1.28)$$

and

$$(\vec{C} \times \vec{D})^i = \epsilon^{ijk} g_{jm} C^m g_{kl} D^l. \quad (1.29)$$

Of course, it is equally reasonable to write

$$(\vec{C} \times \vec{D})^i = g^{ij} \epsilon_{jkl} C^k D^l, \quad \epsilon_{ijk} \stackrel{\text{def}}{=} g_{il} g_{jm} g_{kn} \epsilon^{lmn}. \quad (1.30)$$

A-ha: Cartesian coordinate systems are indeed very simple, in that there $g_{ik} = \delta_{ik}$ form an identity matrix. And, yes—this simplicity is preserved only by constant rotations, *not* by general coordinate transformations!

In n dimensions, the ϵ^{\dots} symbol will have n indices, but will be defined through a formula very much like (1.27). While its utility in writing the cross-product is limited to three dimensions, it can be used to write a perfectly general formula for the determinant of a matrix:

$$\begin{aligned} \det[\mathbf{L}] &= \frac{1}{n!} \epsilon^{i_1 \dots i_n} \epsilon^{j_1 \dots j_n} L_{i_1 j_1} \dots L_{i_n j_n}, \\ &= \epsilon^{i_1 \dots i_n} L_{i_1 1} \dots L_{i_n n}; \end{aligned} \quad (1.31)$$

$$\begin{aligned} \det[\mathbf{M}] &= \frac{1}{n!} \epsilon^{i_1 \dots i_n} \epsilon_{j_1 \dots j_n} M_{i_1}^{j_1} \dots M_{i_n}^{j_n}, \\ &= \epsilon^{i_1 \dots i_n} M_{i_1}^1 \dots M_{i_n}^n : \end{aligned} \quad (1.32)$$

and

$$\begin{aligned}\det[\mathbf{N}] &= \frac{1}{n!} \epsilon_{i_1 \dots i_n} \epsilon_{j_1 \dots j_n} N^{i_1 j_1} \dots N^{i_n j_n} , \\ &= \epsilon_{i_1 \dots i_n} N^{i_1 1} \dots N^{i_n n} .\end{aligned}\tag{1.33}$$

Here, \mathbf{L} is a twice-covariant matrix (rank-two tensor), \mathbf{M} a mixed one and \mathbf{N} is a twice-contravariant matrix.

1.4. Further stuff...

There is another contradiction in the textbook. On p.158, it states that the vectors $\vec{\epsilon}_1, \vec{\epsilon}_2, \vec{\epsilon}_3$ are not necessarily orthogonal. Yet, on the next page, Eq. (3.128), it states that $\vec{\epsilon}_i = h_i \hat{e}_i$, for each i and with no summation. But if each $\vec{\epsilon}_i$ is simply proportional (by a factor h_i) to the \hat{e}_i —*which were treated throughout as orthogonal*—then the $\vec{\epsilon}_i$ must be orthogonal also! Remember that the scaling coefficients h_i were defined as the square-roots of the diagonal elements of the metric, $h_i = \sqrt{g_{ii}}$, and *for orthogonal systems only!* The definition of the *basis vectors*

$$\vec{\epsilon}_i \stackrel{\text{def}}{=} \frac{\partial \vec{r}}{\partial x^i}\tag{1.34}$$

of course makes sense in general and the formulae *after* Eq. (3.128) are correct in general. The paragraph between Eq. (3.127) and Eq. (3.128), including the latter, are correct only for orthogonal systems, since the \hat{e}_i *were always treated as orthogonal*.

By contrast, throughout these notes, the basis vectors \hat{e}_i were never accused of orthogonality or, Heavens forbid, having their length fixed to 1; a set of basis vectors $\{\hat{e}_i\}$ we presume general until proven otherwise!

Finally, note that merely counting the free indices on a quantity does not necessarily tell how that quantity transforms (and this is *not* contradicting the statements above). Facing an unknown quantity, the burden of showing that it *does* transform as a respectable tensor should lie not on the quantity but on you. Consider, for example, transforming the derivative of a vector:

$$\frac{\partial \tilde{A}^j}{\partial \tilde{x}^i} = \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial}{\partial x^k} \left(\frac{\partial \tilde{x}^j}{\partial x^l} A^l \right) = \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial \tilde{x}^j}{\partial x^l} \frac{\partial A^l}{\partial x^k} + \frac{\partial x^k}{\partial \tilde{x}^i} \frac{\partial^2 \tilde{x}^j}{\partial x^l \partial x^k} A^l .\tag{1.35}$$

The second term reveals that $\frac{\partial \tilde{A}^j}{\partial \tilde{x}^i}$ is not a tensor, not even $\frac{\partial A^i}{\partial x^i}$, which we could naïvely think of as the gradient of the contravariant vector \vec{A} .

The remedy for this is to replace the usual partial derivative with another one, the *covariant derivative operator*, but that and its consequences would make these notes considerably longer, which was not the intention. Suffice it merely to note that (ah, yes; rather confusingly) the components of this covariant derivative are denoted by ∇_i . It acts as follows:

$$\nabla_i f = \frac{\partial f}{\partial x^i} ,\tag{1.36a}$$

$$\nabla_i A^j = \frac{\partial A^j}{\partial x^i} + \Gamma_{kj}^j A^k ,\tag{1.36b}$$

$$\nabla_i A_j = \frac{\partial A_j}{\partial x^i} - \Gamma_{ij}^k A_k ,\tag{1.36c}$$

$$\nabla_i A_{jk} = \frac{\partial A_{jk}}{\partial x^i} - \Gamma_{ij}^l A_{lk} - \Gamma_{ik}^l A_{jl} , \quad (1.36d)$$

$$\nabla_i A_j^k = \frac{\partial A_j^k}{\partial x^i} - \Gamma_{ij}^l A_l^k + \Gamma_{il}^k A_j^l , \quad (1.36e)$$

and so on: an additional Γ -term is added per index, positive for superscripts, negative for subscripts. This Γ -object is defined entirely in terms of the metric (and its inverse) as

$$\Gamma_{jk}^i \stackrel{\text{def}}{=} \frac{1}{2} g^{il} \left(\frac{\partial g_{jl}}{\partial x^k} + \frac{\partial g_{lk}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l} \right) . \quad (1.37)$$