A Primer on

Tensor Calculus

David A. Clarke

Saint Mary's University, Halifax NS, Canada

dclarke@ap.smu.ca

June, 2011

Copyright © David A. Clarke, 2011

Contents

Preface

These notes stem from my own need to refresh my memory on the fundamentals of tensor calculus, having seriously considered them last some 25 years ago in grad school. Since then, while I have had ample opportunity to teach, use, and even program numerous ideas from vector calculus, tensor analysis has faded from my consciousness. How much it had faded became clear recently when I tried to program the viscosity tensor into my fluids code, and couldn't account for, much less derive, the myriad of "strange terms" (ultimately from the dreaded "Christ-awful" symbols) that arise when programming a tensor quantity valid in curvilinear coordinates.

My goal here is to reconstruct my understanding of tensor analysis enough to make the connexion between covarient, contravariant, and physical vector components, to understand the usual vector derivative constructs (∇, $\nabla\cdot$, $\nabla\times$) in terms of tensor differentiation, to put *dyads* (*e.g.*, $\nabla\vec{v}$) into proper context, to understand how to derive certain identities involving tensors, and finally, the true test, how to program a realistic viscous tensor to endow a fluid with the non-isotropic stresses associated with Newtonian viscosity in curvilinear coordinates.

Inasmuch as these notes may help others, the reader is free to use, distribute, and modify them as needed so long as they remain in the public domain and are passed on to others free of charge.

David Clarke
Saint Mary's University
June, 2011

Primers by David Clarke:

1. A FORTRAN Primer
2. A UNIX Primer
3. A DBX (debugger) Primer
4. A Primer on Tensor Calculus

I also give a link to David R. Wilkins' excellent primer Getting Started with LaTeX, in which I have added a few sections on adding figures, colour, and HTML links.

A Primer on Tensor Calculus

1 Introduction

In physics, there is an overwhelming need to formulate the basic laws in a so-called *invariant* form; that is, one that does not depend on the chosen coordinate system. As a start, the freshman university physics student learns that in ordinary Cartesian coordinates, Newton's Second Law, $\sum_i \vec{F}_i = m\vec{a}$, has the identical form regardless of which *inertial* frame of reference (not accelerating with respect to the background stars) one chooses. Thus two observers taking independent measures of the forces and accelerations would agree on each measurement made, regardless of how rapidly one observer is moving relative to the other so long as neither observer is accelerating.

However, the sophomore student soon learns that if one chooses to examine Newton's Second Law in a curvilinear coordinate system, such as right-cylindrical or spherical polar coordinates, new terms arise that stem from the fact that the orientation of some coordinate unit vectors change with position. Once these terms, which resemble the *centrifugal* and *Coriolis* terms appearing in a rotating frame of reference, have been properly accounted for, physical laws involving vector quantities can once again be made to "look" the same as they do in Cartesian coordinates, restoring their "invariance".

Alas, once the student reaches their junior year, the complexity of the problems has forced the introduction of rank 2 constructs such as matrices to describe certain physical quantities (*e.g.*, moment of inertia, viscosity, spin) and in the senior year, Riemannian geometry and general relativity require mathematical entities of still higher rank. The tools of vector analysis are simply incapable of allowing one to write down the governing laws in an invariant form, and one has to adopt a different mathematics from the vector analysis taught in the freshman and sophomore years.

Tensor calculus is that mathematics. Clues that tensor-like entities are ultimately needed exist even in a first year physics course. Consider the task of expressing a velocity as a vector quantity. In Cartesian coordinates, the task is rather trivial and no ambiguities arise. Each component of the vector is given by the rate of change of the object's coordinates as a function of time:

$$\vec{v} = (\dot{x}, \dot{y}, \dot{z}) = \dot{x}\,\hat{e}_x + \dot{y}\,\hat{e}_y + \dot{z}\,\hat{e}_z, \tag{1}$$

where I use the standard notation of an "over-dot" for time differentiation, and where $\hat{e}_x$ is the unit vector in the x-direction, *etc.* Each component has the unambiguous units of $\mathrm{m\,s^{-1}}$, the unit vectors point in the same direction no matter where the object may be, and the velocity is completely well-defined.

Ambiguities start when one wishes to express the velocity in spherical-polar coordinates, for example. If, following equation (1), we write the velocity components as the time-derivatives of the coordinates, we might write

$$\vec{v} = (\dot{r}, \dot{\vartheta}, \dot{\varphi}). \tag{2}$$

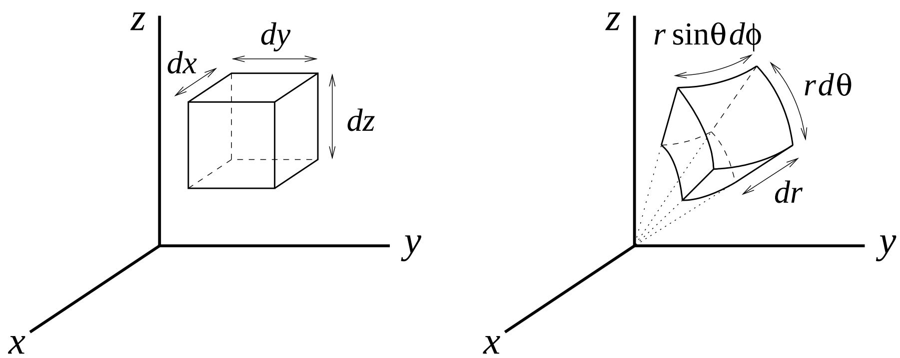

Figure 1: (left) A differential volume in Cartesian coordinates, and (right) a differential volume in spherical polar coordinates, both with their edge-lengths indicated.

An immediate "cause for pause" is that the three components do not share the same "units", and thus we cannot expand this ordered triple into a series involving the respective unit vectors as was done in equation (1). A little reflection might lead us to examine a differential "box" in each of the coordinate systems as shown in Fig. 1. The sides of the Cartesian box have length dx, dy, and dz, while the spherical polar box has sides of length dr, $r\,d\vartheta$, and $r\,\sin\vartheta\,d\varphi$. We might argue that the components of a physical velocity vector should be the lengths of the differential box divided by dt, and thus:

$$\vec{v} \,=\, (\dot{r}, r\,\dot{\vartheta}, r\sin\vartheta\,\dot{\varphi}) \,=\, \dot{r}\,\hat{e}_r + r\,\dot{\vartheta}\,\hat{e}_\vartheta + r\sin\vartheta\,\dot{\varphi}\,\hat{e}_\varphi, \tag{3}$$

which addresses the concern about units. So which is the "correct" form?

In the pages that follow, we shall see that a *tensor* may be designated as *contravariant*, *covariant*, or *mixed*, and that the velocity expressed in equation (2) is in its contravariant form. The velocity vector in equation (3) corresponds to neither the covariant nor contravariant form, but is in its so-called *physical* form that we would measure with a speedometer. Each form has a purpose, no form is any more fundamental than the other, and all are linked via a very fundamental tensor called the *metric*. Understanding the role of the metric in linking the various forms of tensors[1] and, more importantly, in differentiating tensors is the basis of *tensor calculus*, and the subject of this primer.

[1]Examples of tensors the reader is already familiar with include scalars (rank 0 tensors) and vectors (rank 1 tensors).

2 Definition of a tensor

As mentioned, the need for a mathematical construct such as tensors stems from the need to know how the functional dependence of a physical quantity on the position coordinates changes with a change in coordinates. Further, we wish to render the fundamental laws of physics relating these quantities *invariant* under coordinate transformations. Thus, while the functional form of the acceleration vector may change from one coordinate system to another, the functional changes to $\vec{F}$ and m will be such that $\vec{F}$ will always be equal to $m\vec{a}$, and not some other function of m, $\vec{a}$, and/or some other variables or constants depending on the coordinate system chosen.

Consider two coordinate systems, x_i and $\tilde{x}_i$, in an n-dimensional space where $i = 1, 2, \dots, n$[2]. x_i and $\tilde{x}_i$ could be two Cartesian coordinate systems, one moving at a constant velocity relative to the other, or x_i could be Cartesian coordinates and $\tilde{x}_i$ spherical polar coordinates whose origins are coincident and in relative rest. Regardless, one should be able, in principle, to write down the coordinate transformations in the following form:

$$\tilde{x}_i = \tilde{x}_i(x_1, x_2, \dots, x_n), \tag{4}$$

one for each i, and their inverse transformations:

$$x_i = x_i(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n). \tag{5}$$

Note that which of equations (4) and (5) is referred to as the "transformation", and which as the "inverse" is completely arbitrary. Thus, in the first example where the Cartesian coordinate system $\tilde{x}_i = (\tilde{x}, \tilde{y}, \tilde{z})$ is moving with velocity v along the $+x$ axis of the Cartesian coordinate system $x_i = (x, y, z)$, the transformation relations and their inverses are:

$$\left.\begin{aligned} \tilde{x} &= x - vt, & x &= \tilde{x} + vt, \\ \tilde{y} &= y, & y &= \tilde{y}, \\ \tilde{z} &= z, & z &= \tilde{z}. \end{aligned}\right\} \tag{6}$$

For the second example, the coordinate transformations and their inverses between Cartesian, $x_i = (x, y, z)$, and spherical polar, $\tilde{x}_i = (r, \vartheta, \varphi)$ coordinates are:

$$\left.\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, & x &= r \sin\vartheta \cos\varphi, \\ \vartheta &= \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{z}\right), & y &= r \sin\vartheta \sin\varphi, \\ \varphi &= \tan^{-1}\left(\frac{y}{x}\right), & z &= r \cos\vartheta. \end{aligned}\right\} \tag{7}$$

Now, let f be some function of the coordinates that represents a physical quantity of interest. Consider again two generic coordinate systems, x_i and $\tilde{x}_i$, and assume their transformation relations, equations (4) and (5), are known. If the components of the gradient

[2]In physics, n is normally 3 or 4 depending on whether the discussion is non-relativistic or relativistic, though our discussion matters little on a specific value of n. Only when we are speaking of the curl and cross-products in general will we deliberately restrict our discussion to 3-space.

of f in x_j, namely $\partial f/\partial x_j$, are known, then we can find the components of the gradient in $\tilde{x}_i$, namely $\partial f/\partial \tilde{x}_i$, by the chain rule:

$$\frac{\partial f}{\partial \tilde{x}_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial \tilde{x}_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial \tilde{x}_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial \tilde{x}_i} = \sum_{j=1}^{n} \frac{\partial x_j}{\partial \tilde{x}_i}\frac{\partial f}{\partial x_j}. \tag{8}$$

Note that the coordinate transformation information appears as partial derivatives of the *old* coordinates, x_j, with respect to the *new* coordinates, $\tilde{x}_i$.

Next, let us ask how a differential of one of the new coordinates, $d\tilde{x}_i$, is related to differentials of the old coordinates, dx_i. Again, an invocation of the chain rule yields:

$$d\tilde{x}_i = dx_1\frac{\partial \tilde{x}_i}{\partial x_1} + dx_2\frac{\partial \tilde{x}_i}{\partial x_2} + \cdots + dx_n\frac{\partial \tilde{x}_i}{\partial x_n} = \sum_{j=1}^{n} \frac{\partial \tilde{x}_i}{\partial x_j}\, dx_j. \tag{9}$$

This time, the coordinate transformation information appears as partial derivatives of the *new* coordinates, $\tilde{x}_i$, with respect to the *old* coordinates, x_j, and the inverse of equation (8).

We now redefine what it means to be a *vector* (equally, a *rank 1 tensor*).

Definition 2.1. The components of a *covariant vector* transform like a gradient and obey the transformation law:

$$\tilde{A}_i = \sum_{j=1}^{n} \frac{\partial x^j}{\partial \tilde{x}^i} A_j. \tag{10}$$

Definition 2.2. The components of a *contravariant vector* transform like a coordinate differential and obey the transformation law:

$$\tilde{A}^i = \sum_{j=1}^{n} \frac{\partial \tilde{x}^i}{\partial x^j} A^j. \tag{11}$$

It is customary, as illustrated in equations (10) and (11), to leave the indices of covariant tensors as subscripts, and to raise the indices of contravariant tensors to superscripts: "co-low, contra-high"[3]. In this convention, $dx_i \to dx^i$. As a practical modification to this rule, because of the difference between the definitions of covariant and contravariant components (equations 10 and 11), a *contravariant* index in the denominator is equivalent to a *covarient* index in the numerator, and *vice versa*. Thus, in the construct $\partial x^j/\partial \tilde{x}^i$, j is contravariant while i is considered to be covariant.

Superscripts indicating raising a variable to some power will generally be clear by context, but where there is any ambiguity, indices representing powers will be enclosed in square brackets. Thus, A^2 will normally be, from now on, the "2-component of the contravariant vector A", whereas $A^{[2]}$ will be "A-squared" when A^2 could be ambiguous.

[3]Thanks to Rob Thacker, SMU, for this handy mnemonic.

Finally, we shall adopt here, as is done most everywhere else, the *Einstein summation convention* in which a covariant index followed by the identical contravariant index (or *vice versa*) is implicitly summed over the index *without the use of a summation sign*, rendering the repeated index a dummy index. On rare occasions where a sum is to be taken over two repeated covariant or two repeated contravariant indices, a summation sign will be given explicitly. Conversely, if properly repeated indices (*e.g.*, one covariant, one contravariant) are *not* to be summed, a note to that effect will be given. Further, any indices enclosed in parentheses [*e.g.*, (i)] will not be summed. Thus, A_iB^i is normally summed while A_iB_i, A^iB^i, and $A_{(i)}B^{(i)}$ are not.

To the uninitiated who may think at first blush that this convention may be fraught with exceptions, it turns out to be remarkably robust and rarely will it pose any ambiguities. In tensor analysis, it is rare that two properly repeated indices should not, in fact, be summed. It is equally rare that two repeated covariant (or contravariant) indices should be summed, and rarer still that an index appears more than twice in any given term.

As a first illustration, applying the Einstein summation convention changes equations (10) and (11) to:

$$\tilde{A}_i = \frac{\partial x^j}{\partial \tilde{x}^i} A_j, \qquad \text{and} \qquad \tilde{A}^i = \frac{\partial \tilde{x}^i}{\partial x^j} A^j,$$

respectively, where summation is implicit over the index j in both cases.

Remark 2.1. While dx^i is the prototypical rank 1 contravariant tensor (*e.g.*, equation 9), x^i is not a tensor as its transformation follows neither equations (10) nor (11). Still, we will follow the up-down convention for coordinate indices as it serves a purpose to distinguish between covariant-*like* and contravariant-*like* coordinates. It will usually be the case anyway that x^i will appear as part of dx^i or $\partial/\partial x^i$.

Tensors of higher rank[4] are defined in an entirely analogous way. A tensor of dimension m (each index varies from 1 to m) and rank n (number of indices) is an entity that, under an arbitrary coordinate transformation, transforms as:

$$\boxed{\tilde{T}_{i_1\dots i_p}{}^{k_1\dots k_q} = \frac{\partial x^{j_1}}{\partial \tilde{x}^{i_1}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}} T_{j_1\dots j_p}{}^{l_1\dots l_q},} \tag{12}$$

where $p + q = n$, and where the indices $i_1, \dots, i_p$ and $j_1, \dots, j_p$ are covariant indices and $k_1, \dots, k_q$ and $l_1, \dots, l_q$ are contravariant indices. Indices that appear just once in a term (*e.g.*, $i_1, \dots, i_p$ and $k_1, \dots, k_q$ in equation 12) are called *free* indices, while indices appearing twice—one covariant and one contravariant—(*e.g.*, $j_1, \dots, j_p$ and $l_1, \dots, l_q$ in equation 12), are called *dummy* indices as they disappear after the implied sum is carried forth. *In a*

[4]There is great potential for confusion on the use of the term *rank*, as it is not used consistently in the literature. In the context of matrices, if the n column vectors in an $m \times n$ matrix (with "tensor-rank" 2) can all be expressed as a linear combination of $r \le \min(m, n)$ m-dimensional vectors, that matrix has a "matrix-rank" r which, of course, need not be 2. For this reason, some authors prefer to use *order* rather than *rank* for tensors so that a scalar is an *order-0* tensor, a vector an *order-1* tensor, and a matrix an *order-2* tensor. Still other authors use *dimension* instead of *rank*, although this then gets confused with the *dimension* of a vector (number of linearly independent vectors that span the parent vector space).

Despite its potential for confusion, I use the term *rank* for the number of indices on a tensor, which is in keeping with the most common practise.

valid tensor relationship, each term, whether on the left or right side of the equation, must have the same free indices each in the same position. If a certain free index is covariant (contravariant) in one term, it must be covariant (contravariant) in all terms.

If $q = 0$ ($p = 0$), then all indices are covariant (contravariant) and the tensor is said to be covariant (contravariant). Otherwise, if the tensor has both covariant and contravariant indices, it is said to be *mixed*. In general, the order of the indices is important, and we deliberately write the tensor as $T_{j_1 \dots j_p}{}^{l_1 \dots l_q}$, and not $T^{l_1 \dots l_q}_{j_1 \dots j_p}$. However, there is no reason to expect all contravariant indices to follow the covariant indices, nor for all covariant indices to be listed contiguously. Thus and for example, one could have $T_i{}^j{}_{kl}{}^m$ if, indeed, the first, third, and fourth indices were covariant, and the second and fifth indices were contravariant.

Remark 2.2. Rank 2 tensors of dimension m can be represented by $m \times m$ square matrices. A matrix that is an element of a vector space is a rank 2 tensor. Rank 3 tensors of dimension m would be represented by an $m \times m \times m$ cube of values, *etc.*

Remark 2.3. In traditional vector analysis, one is forever moving back and forth between considering vectors as a whole (*e.g.*, $\vec{v}$), or in terms of its components relative to some coordinate system (*e.g.*, v_x). This, then, leads one to worry whether a given relationship is true for all coordinate systems (*e.g.*, vector "identities" such as: $\nabla \cdot f\vec{A} = f\nabla \cdot \vec{A} + \vec{A} \cdot \nabla f$), or whether it is true only in certain coordinate systems [*e.g.*, $\nabla \cdot (\vec{A}\vec{B}) = (\nabla \cdot B_x\vec{A}, \nabla \cdot B_y\vec{A}, \nabla \cdot B_z\vec{A})$ is true in Cartesian coordinates only]. The formalism of tensor analysis eliminates both of these concerns by writing everything down in terms of a "typical tensor component" where all "geometric factors", which have yet to be discussed, have been safely accounted for in the notation. As such, all equations are written in terms of tensor components, and rarely is a tensor written down without its indices. As we shall see, this both simplifies the notation and renders unambiguous the invariance of certain relationships under arbitrary coordinate transformations.

In the remainder of this section, we make a few definitions and prove a few theorems that will be useful throughout the rest of this primer.

Theorem 2.1. *The sum (or difference) of two like-tensors is a tensor of the same type.*

Proof. This is a simple application of equation (12). Consider two like-tensors (*i.e.*, identical indices), S and T, each transforming according to equation (12). Adding the LHS and the RHS of these transformation equations (and defining $\mathsf{R} = \mathsf{S} + \mathsf{T}$), one gets:

$$\begin{aligned}
\tilde{R}_{i_1 \dots i_p}{}^{k_1 \dots k_q} &\equiv \tilde{S}_{i_1 \dots i_p}{}^{k_1 \dots k_q} + \tilde{T}_{i_1 \dots i_p}{}^{k_1 \dots k_q} \\
&= \frac{\partial x^{j_1}}{\partial \tilde{x}^{i_1}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}} S_{j_1 \dots j_p}{}^{l_1 \dots l_q} \\
&+ \frac{\partial x^{j_1}}{\partial \tilde{x}^{i_1}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}} T_{j_1 \dots j_p}{}^{l_1 \dots l_q} \\
&= \frac{\partial x^{j_1}}{\partial \tilde{x}^{i_1}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}} (S_{j_1 \dots j_p}{}^{l_1 \dots l_q} + T_{j_1 \dots j_p}{}^{l_1 \dots l_q}) \\
&= \frac{\partial x^{j_1}}{\partial \tilde{x}^{i_1}} \cdots \frac{\partial x^{j_p}}{\partial \tilde{x}^{i_p}} \frac{\partial \tilde{x}^{k_1}}{\partial x^{l_1}} \cdots \frac{\partial \tilde{x}^{k_q}}{\partial x^{l_q}} R_{j_1 \dots j_p}{}^{l_1 \dots l_q}.
\end{aligned}$$

□

Definition 2.3. A rank 2 *dyad*, D, results from taking the *dyadic* product of two vectors (rank 1 tensors), $\vec{A}$ and $\vec{B}$, as follows:

$$D_{ij} = A_i B_j, \quad D_i{}^j = A_i B^j, \quad D^i_j = A^i B_j, \quad D^{ij} = A^i B^j, \tag{13}$$

where the ij^{th} component of D, namely $A_i B_j$, is just the ordinary product of the i^{th} element of $\vec{A}$ with the j^{th} element of $\vec{B}$.

The dyadic product of two covariant (contravariant) vectors yields a covariant (contravariant) dyad (first and fourth of equations 13), while the dyadic product of a covariant vector and a contravariant vector yields a *mixed* dyad (second and third of equations 13). Indeed, dyadic products of three or more vectors can be taken to create a dyad of rank 3 or higher (*e.g.*, $D_i{}^j{}_k = A_i B^j C_k$, etc).

Theorem 2.2. *A rank 2 dyad is a rank 2 tensor.*

Proof. We need only show that a rank 2 dyad transforms as equation (12). Consider a mixed dyad, $\tilde{D}_k{}^l = \tilde{A}_k \tilde{B}^l$, in a coordinate system $\tilde{x}^l$. Since we know how the vectors transform to a different coordinate system, x^i, we can write:

$$\tilde{D}_k{}^l = \tilde{A}_k \tilde{B}^l = \left(\frac{\partial x^i}{\partial \tilde{x}^k} A_i\right)\left(\frac{\partial \tilde{x}^l}{\partial x^j} B^j\right) = \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial \tilde{x}^l}{\partial x^j} A_i B^j = \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial \tilde{x}^l}{\partial x^j} D_i{}^j.$$

Thus, from equation (12), D transforms as a mixed tensor of rank 2. A similar argument can be made for a purely covariant (contravariant) dyad of rank 2 and, by extension, of arbitrary rank. □

Remark 2.4. By dealing with only a typical component of a tensor (and thus a real or complex number), all arithmetic is ordinary multiplication and addition, and everything commutes: $A_i B^j = B^j A_i$, for example. Conversely, in vector and matrix algebra when one is dealing with the entire vector or matrix, multiplication does not follow the usual rules of scalar multiplication and, in particular, is not commutative. In many ways, this renders tensor algebra much simpler than vector and matrix algebra.

Definition 2.4. If A_{ij} is a rank 2 covariant tensor and B^{kl} is a rank 2 contravariant tensor, then they are each other's *inverse* if:

$$A_{ij} B^{jk} = \delta_i{}^k,$$

where $\delta_i{}^k = 1$ if $i = k$, 0 otherwise is the usual *Kronecker delta*.

In a similar vein, $A_i{}^j B_j{}^k = \delta_i{}^k$ and $A^i{}_j B^j{}_k = \delta^i{}_k$ are examples of inverses for mixed rank 2 tensors. One can even have "inverses" of rank 1 tensors: $\boldsymbol{e}_i \boldsymbol{e}^j = \delta_i{}^j$, though this property is usually referred to as *orthogonality*.

Note that the concepts of invertibility and orthogonality take the place of "division" in tensor algebra. Thus, one would never see a tensor element in the denominator of a fraction and something like $C_i{}^k = A_{ij}/B_{jk}$ is *never* written. Instead, one would write $C_i{}^k B_{jk} = A_{ij}$ and, if it were critical that C be isolated, one would write $C_i{}^k = A_{ij}(B_{jk})^{-1} = A_{ij} D^{jk}$ if

D^{jk} were, in fact, the inverse of B_{jk}. A tensor element could appear in the numerator of a fraction where the demonimator is a scalar (*e.g.*, $A_{ij}/2$) or a *physical component* of a vector (as introduced in the next section), but never in the denominator.

I note in haste that while a derivative, $\partial x^i/\partial\tilde{x}^j$, may look like an exception to this rule, it is a notational exception only. In taking a derivative, one is not really taking a fraction. And while dx^i is a tensor, Δx^i is not and thus the actual fraction $\Delta x^i/\Delta\tilde{x}^j$ is allowed in tensor algebra since the denominator is not a tensor element.

Theorem 2.3. *The derivatives $\partial x^i/\partial\tilde{x}^k$ and $\partial\tilde{x}^k/\partial x^j$ are each other's inverse. That is,*

$$\frac{\partial x^i}{\partial\tilde{x}^k}\frac{\partial\tilde{x}^k}{\partial x^j} = \delta^i{}_j.$$

Proof. This is a simple application of the chain rule. Thus,

$$\frac{\partial x^i}{\partial\tilde{x}^k}\frac{\partial\tilde{x}^k}{\partial x^j} = \frac{\partial x^i}{\partial x^j} = \delta^i{}_j, \tag{14}$$

where the last equality is true by virtue of the independence of the coordinates x^i. □

Remark 2.5. If one considers $\partial x^i/\partial\tilde{x}^k$ and $\partial\tilde{x}^k/\partial x^j$ as, respectively, the $(i,k)^{\text{th}}$ and $(k,j)^{\text{th}}$ elements of $m \times m$ matrices then, with the implied summation, the LHS of equation (14) is simply following ordinary matrix multiplication, while the RHS is the $(i,j)^{\text{th}}$ element of the identity matrix. It is in this way that $\partial x^i/\partial\tilde{x}^k$ and $\partial\tilde{x}^k/\partial x^j$ are each other's inverse.

Definition 2.5. A *tensor contraction* occurs when one of a tensor's free covarient indices is set equal to one of its free contravariant indices. In this case, a sum is performed on the now repeated indices, and the result is a tensor with two fewer free indices.

Thus, and for example, $T_{ij}{}^j$ is a contraction on the second and third indices of the rank 3 tensor $T_{ij}{}^k$. Once the sum is performed over the repeated indices, the result is a rank 1 tensor (vector). Thus, if we use T to designate the contracted tensor as well (something we are not obliged to do, but certainly may), we would write:

$$T_{ij}{}^j = T_i.$$

Remark 2.6. Contractions are only ever done between one covariant index and one contravariant index, never between two covariant indices nor two contravariant indices.

Theorem 2.4. *A contraction of a rank 2 tensor (its* trace*) is a scalar whose value is independent of the coordinate system chosen. Such a scalar is referred to as a* rank 0 tensor*.*

Proof. Let $T = T_i{}^i$ be the trace of the tensor, T. If $\tilde{T}_k{}^l$ is a tensor in coordinate system $\tilde{x}^k$, then its trace transforms to coordinate system x^i according to:

$$\tilde{T}_k{}^k = \frac{\partial x^i}{\partial\tilde{x}^k}\frac{\partial\tilde{x}^k}{\partial x^j}T_i{}^j = \delta^i{}_j T_i{}^j = T_i{}^i = T.$$ □

It is important to note the role played by the fact that $\tilde{T}_k{}^l$ is a tensor, and how this gave rise to the Kronecker delta (Theorem 2.3) which was needed in proving the invariance of the trace (*i.e.*, that the trace has the same value regardless of coordinate system).

3 The metric

In an arbitrary m-dimensional coordinate system, x^i, the differential displacement vector is:

$$d\vec{r} = (h_{(1)}dx^1, h_{(2)}dx^2, \ldots, h_{(m)}dx^m) = \sum_{i=1}^{m} h_{(i)}dx^i\hat{e}_{(i)}, \tag{15}$$

where $\hat{e}_{(i)}$ are the physical (not covariant) unit vectors, and where $h_{(i)} = h_{(i)}(x^1, \ldots, x^m)$ are scale factors (not tensors) that depend, in general, on the coordinates and endow each component with the appropriate units of length. The subscript on the unit vector is enclosed in parentheses since it is a vector label (to distinguish it from the other unit vectors spanning the vector space), and not an indicator of a component of a covariant tensor. Subscripts on h are enclosed in parentheses since they, too, do not indicate components of a covariant tensor. In both cases, the parentheses prevent them from triggering an application of the Einstein summation convention should they be repeated. For the three most common orthogonal coordinate systems, the coordinates, unit vectors, and scale factors are:

system	x^i	$\hat{e}_{(i)}$	$(h_{(1)}, h_{(2)}, h_{(3)})$
Cartesian	(x, y, z)	$\hat{e}_x, \hat{e}_y, \hat{e}_z$	$(1, 1, 1)$
cylindrical	(z, ϱ, φ)	$\hat{e}_z, \hat{e}_\varrho, \hat{e}_\varphi$	$(1, 1, \varrho)$
spherical polar	(r, ϑ, φ)	$\hat{e}_r, \hat{e}_\vartheta, \hat{e}_\varphi$	$(1, r, r\sin\vartheta)$

Table 1: Nomenclature for the most common coordinate systems.

In "vector-speak", the length of the vector $d\vec{r}$, given by equation (15), is obtained by taking the "dot product" of $d\vec{r}$ with itself. Thus,

$$(dr)^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} h_{(i)}h_{(j)}\hat{e}_{(i)} \cdot \hat{e}_{(j)}dx^i dx^j, \tag{16}$$

where $\hat{e}_{(i)} \cdot \hat{e}_{(j)} \equiv \cos\theta_{(ij)}$ are the *directional cosines* which are 1 when $i = j$. For orthogonal coordinate systems, $\cos\theta_{(ij)} = 0$ for $i \neq j$, thus eliminating the "cross terms". For non-orthoginal systems, the off-diagonal directional cosines are not, in general, zero and the cross-terms remain.

Definition 3.1. The *metric*, g_{ij}, is given by:

$$g_{ij} = h_{(i)}h_{(j)}\hat{e}_{(i)} \cdot \hat{e}_{(j)}, \tag{17}$$

which, by inspection, is symmetric under the interchange of its indices; $g_{ij} = g_{ji}$.

Remark 3.1. For an orthogonal coordinate system, the metric is given by: $g_{ij} = h_{(i)}h_{(j)}\delta_{ij}$, which reduces further to δ_{ij} for Cartesian coordinates.

Thus equation (16) becomes:

$$(dr)^2 = g_{ij}dx^i dx^j, \tag{18}$$

where the summation on i and j is now implicit, presupposing the following theorem:

Theorem 3.1. *The metric is a rank 2 covariant tensor.*

Proof. Because $(dr)^2$ is the square of a distance between two physical points, it must be invariant under coordinate transformations. Thus, consider $(dr)^2$ in the coordinate systems $\tilde{x}^k$ and x^i:

$$(dr)^2 = \tilde{g}_{kl}\, d\tilde{x}^k d\tilde{x}^l = g_{ij}\, dx^i dx^j = g_{ij} \frac{\partial x^i}{\partial \tilde{x}^k} d\tilde{x}^k \frac{\partial x^j}{\partial \tilde{x}^l} d\tilde{x}^l = \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^l}\, g_{ij}\, d\tilde{x}^k d\tilde{x}^l$$

$$\Rightarrow \quad \left(\tilde{g}_{kl} - \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^l}\, g_{ij}\right) d\tilde{x}^k d\tilde{x}^l = 0,$$

which must be true $\forall\ d\tilde{x}^k d\tilde{x}^l$. This can be true only if,

$$\tilde{g}_{kl} = \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^l}\, g_{ij}, \tag{19}$$

and, by equation (12), g_{ij} transforms as a rank 2 covariant tensor. □

Definition 3.2. The *conjugate metric*, g^{kl}, is the inverse to the metric tensor, and therefore satisfies:

$$g^{kp} g_{ip} = g_{ip} g^{kp} = \delta_i^{\ k}. \tag{20}$$

It is left as an exercise to show that the conjugate metric is a rank 2 contravariant tensor. (Hint: use the invariance of the Kronecker delta.)

Definition 3.3. A *conjugate tensor* is the result of multiplying a tensor with the metric, then contracting one of the indices of the metric with one of the indices of the tensor.

Thus, two examples of conjugates for the rank n tensor $T_{i_1\dots i_p}{}^{j_1\dots j_q}$, $p+q=n$, include:

$$T_{i_1\dots i_{r-1}}{}^{k}{}_{i_{r+1}\dots i_p}{}^{j_1\dots j_q} = g^{k i_r} T_{i_1\dots i_p}{}^{j_1\dots j_q}, \quad 1 \le r \le p; \tag{21}$$

$$T_{i_1\dots i_p}{}^{j_1\dots j_{s-1}}{}_{l}{}^{j_{s+1}\dots j_q} = g_{l j_s} T_{i_1\dots i_p}{}^{j_1\dots j_q}, \quad 1 \le s \le q. \tag{22}$$

An operation like equation (21) is known as *raising an index* (covariant index i_r is replaced with contravariant index k) while equation (22) is known as *lowering an index* (contravariant index j_s is replaced with covariant index l). For a tensor with p covariant and q contravariant indices, one could write down p conjugate tensors with a single index raised and q conjugate tensors with a single index lowered. Contracting a tensor with the metric several times will raise or lower several indices, each representing a conjugate tensor to the original. Associated with every rank n tensor are 2^{n-1} conjugate tensors all with rank n.

The attempt to write equations (21) and (22) for a general rank n tensor has made them rather opaque, so it is useful to examine the simpler and special case of raising and lowering the index of a rank 1 tensor. Thus,

$$A^j = g^{ij} A_i; \qquad A_i = g_{ij} A^j. \tag{23}$$

Rank 2 tensors can be similarly examined. As a first example, we define the covariant coordinate differential, dx_i, to be:

$$dx^j = g^{ij} dx_i; \qquad dx_i = g_{ij} dx^j.$$

We can find convenient expressions for the metric components of any coordinate system, x^i, if we know how x^i and Cartesian coordinates depend on each other. Thus, if χ^k represent the Cartesian coordinates (x, y, z), and if we know $\chi^k = \chi^k(x^i)$, then using Remark 3.1 we can write:

$$(dr)^2 = \delta_{kl} d\chi^k d\chi^l = \delta_{kl}\left(\frac{\partial \chi^k}{\partial x^i}dx^i\right)\left(\frac{\partial \chi^l}{\partial x^j}dx^j\right) = \left(\delta_{kl}\frac{\partial \chi^k}{\partial x^i}\frac{\partial \chi^l}{\partial x^j}\right)dx^i dx^j.$$

Therefore, by equations (18) and (20), the metric and its inverse are given by:

$$g_{ij} = \delta_{kl}\frac{\partial \chi^k}{\partial x^i}\frac{\partial \chi^l}{\partial x^j} \quad \text{and} \quad g^{ij} = \delta^{kl}\frac{\partial x^i}{\partial \chi^k}\frac{\partial x^j}{\partial \chi^l}. \tag{24}$$

3.1 Physical components and basis vectors

Consider an m-dimensional space, $\mathbb{R}^m$, spanned by an arbitrary basis of *unit* vectors (not necessarily orthogonal), $\hat{e}_{(i)}$, $i = 1, 2, \ldots, m$. A theorem of first-year linear algebra states that for every $\vec{A} \in \mathbb{R}^m$, there is a unique set of numbers, $A_{(i)}$, such that:

$$\vec{A} = \sum_i A_{(i)}\hat{e}_{(i)}. \tag{25}$$

Definition 3.4. The values $A_{(i)}$ in equation (25) are the *physical components* of $\vec{A}$ relative to the basis set $\hat{e}_{(i)}$.

A physical component of a vector field has the same units as the field. Thus, a physical component of velocity has units $\mathrm{m\,s^{-1}}$, electric field $\mathrm{V\,m^{-1}}$, force N, *etc.* As it turns out, a physical component is neither covariant nor contravariant, and thus the subscripts of physical components are surrounded with parentheses lest they trigger an unwanted application of the summation convention which only applies to a covariant-contravariant index pair. As will be shown in this subsection, all three types of vector components are distinct yet related.

It is a straight-forward, if not tedious, task to find the physical components of a given vector, $\vec{A}$, with respect to a given basis set, $\hat{e}_{(i)}$[5]. Suppose $\vec{A}_{\mathrm{C}}$ and $\hat{e}_{i,\mathrm{C}}$ are "m-tuples" of the components of vectors $\vec{A}$ and $\hat{e}_{(i)}$ relative to "Cartesian-like[6]" coordinates (or any coordinate system, for that matter). To find the m-tuple, $\vec{A}_x$, of the components of $\vec{A}$ relative to the new basis, $\hat{e}_{(i)}$, one does the following calculation:

$$\left[\begin{array}{cccc|c} \hat{e}_{1,\mathrm{C}} & \hat{e}_{2,\mathrm{C}} & \ldots & \hat{e}_{m,\mathrm{C}} & \vec{A}_{\mathrm{C}} \\ \downarrow & \downarrow & \ldots & \downarrow & \downarrow \\ & & & & \end{array}\right] \xrightarrow[\text{row reduce}]{} \left[\begin{array}{c|c} & \vec{A}_x \\ I & \downarrow \\ & \end{array}\right]. \tag{26}$$

The i^{th} element of the m-tuple $\vec{A}_x$ will be $A_{(i)}$, the i^{th} component of $\vec{A}$ relative to $\hat{e}_{(i)}$, and the coefficient for equation (25). Should $\hat{e}_{(i)}$ form an orthogonal basis set, the problem of finding

[5] See, for example, §2.7 of Bradley's *A primer of Linear Algebra*; ISBN 0-13-700328-5

[6] I use the qualifier "like" since Cartesian coordinates are, strictly speaking, 3-D.

$A_{(i)}$ is much simpler. Taking the dot product of equation (25) with $\hat{e}_{(j)}$, then replacing the surviving index j with i, one gets:

$$A_{(i)} = \vec{A} \cdot \hat{e}_{(i)}, \tag{27}$$

where, as a matter of practicality, one may perform the dot product using the Cartesian-like components of the vectors, $\vec{A}_{\rm C}$ and $\hat{e}_{i,{\rm C}}$. It must be emphasised that what many readers may think as the general expression, equation (27), works *only* when $\hat{e}_{(i)}$ form an orthogonal set. Otherwise, equation (26) must be used.

Equation (15) expresses the prototypical vector, $d\vec{r}$, in terms of the unit (physical) basis vectors, $\hat{e}_{(i)}$, and its contravariant components, dx^i [the naked differentials of the coordinates, such as $(dr, d\vartheta, d\varphi)$ in spherical polar coordinates]. Thus, by analogy, we write for any vector $\vec{A}$:

$$\vec{A} = \sum_{i=1}^{m} h_{(i)} A^i \hat{e}_{(i)}. \tag{28}$$

On comparing equations (25) and (28), and given the uniqueness of the physical components, we can immediately write down a relation between the physical and contravariant components:

$$A_{(i)} = h_{(i)} A^i; \qquad A^i = \frac{1}{h_{(i)}} A_{(i)}, \tag{29}$$

where, as a reminder, there is no sum on i. By substituting $A^i = g^{ij} A_j$ in the first of equations (29) and multiplying through the second equation by g_{ij} (and thus triggering a sum on i on both sides of the equation), we get the relationship between the physical and covariant components:

$$A_{(i)} = h_{(i)} g^{ij} A_j; \qquad A_j = \sum_i \frac{g_{ij}}{h_{(i)}} A_{(i)}. \tag{30}$$

For orthogonal coordinates, $g^{ij} = \delta^{ij}/h_{(i)} h_{(j)}$, $g_{ij} = \delta_{ij} h_{(i)} h_{(j)}$, and equations (30) reduce to:

$$A_{(i)} = \frac{1}{h_{(i)}} A_i; \qquad A_j = h_{(j)} A_{(j)}. \tag{31}$$

By analogy, we can also write down the relationships between physical components of higher rank tensors and their contravariant, mixed, and covariant forms. For rank 2 tensors, these are:

$$T_{(ij)} = h_{(i)} h_{(j)} T^{ij} = h_{(i)} h_{(j)} g^{ik} T_k{}^j = h_{(i)} h_{(j)} g^{jl} T^i{}_l = h_{(i)} h_{(j)} g^{ik} g^{jl} T_{kl}, \tag{32}$$

(sums on k and l only) which, for orthogonal coordinates, reduce to:

$$T_{(ij)} = h_{(i)} h_{(j)} T^{ij} = \frac{h_{(j)}}{h_{(i)}} T_i{}^j = \frac{h_{(i)}}{h_{(j)}} T^i{}_j = \frac{1}{h_{(i)} h_{(j)}} T_{ij}. \tag{33}$$

Just as there are covariant, physical, and contravariant tensor components, there are also covariant, physical (*e.g.*, unit), and contravariant basis vectors.

Definition 3.5. Let $\vec{r}_x$ be a displacement vector whose components are expressed in terms of the coordinate system x^i. Then the *covariant basis vector*, $\boldsymbol{e}_i$, is defined to be:

$$\boldsymbol{e}_i \equiv \frac{d\vec{r}_x}{dx^i}.$$

Remark 3.2. Just like the unit basis vector $\hat{e}_{(i)}$, the index on $\boldsymbol{e}_i$ serves to distinguish one basis vector from another, and does not represent a single tensor element as, for example, the subscript in the covariant vector A_i does.

It is easy to see that $\boldsymbol{e}_i$ is a covariant vector, by considering its transformation to a new coordinate system, $\tilde{x}^j$:

$$\tilde{\boldsymbol{e}}_j = \frac{d\vec{r}_{\tilde{x}}}{d\tilde{x}^j} = \frac{\partial x^i}{\partial \tilde{x}^j}\frac{d\vec{r}_x}{dx^i} = \frac{\partial x^i}{\partial \tilde{x}^j}\boldsymbol{e}_i,$$

confirming its covariant character.

From equation (15) and the fact that the x^i are linearly independent, it follows from definition 3.5 that:

$$\boldsymbol{e}_i = h_{(i)}\hat{e}_{(i)}; \qquad \hat{e}_{(i)} = \frac{1}{h_{(i)}}\boldsymbol{e}_i. \tag{34}$$

The contravariant basis vector, $\boldsymbol{e}^j$, is obtained by multiplying the first of equations (34) by g^{ij} (triggering a sum over i on the LHS, and thus on the RHS as well), and replacing $\boldsymbol{e}_i$ in the second with $g_{ij}\boldsymbol{e}^j$:

$$g^{ij}\boldsymbol{e}_i = \boldsymbol{e}^j = \sum_i g^{ij}h_{(i)}\hat{e}_{(i)}; \qquad \hat{e}_{(i)} = \frac{g_{ij}}{h_{(i)}}\boldsymbol{e}^j \tag{35}$$

Remark 3.3. Only the physical basis vectors are actually *unit* vectors and thus unitless, and therefore only they are designated by a "hat" ($\hat{e}$). The covariant basis vector, $\boldsymbol{e}_i$, has units $h_{(i)}$ while the contravariant basis vector, $\boldsymbol{e}^j$, has units $1/h_{(i)}$ and are designated in bold-italic ($\boldsymbol{e}$).

Theorem 3.2. *Regardless of whether the coordinate system is orthogonal,* $\boldsymbol{e}_i \cdot \boldsymbol{e}^j = \delta_i^{\ j}$.

Proof.

$$\boldsymbol{e}_i \cdot \boldsymbol{e}^j = h_{(i)}\hat{e}_{(i)} \cdot \sum_k g^{kj}h_{(k)}\hat{e}_{(k)} = \sum_k g^{kj}\underbrace{h_{(i)}h_{(k)}\hat{e}_{(i)} \cdot \hat{e}_{(k)}}_{g_{ik}\ (\text{eq}^{\text{n}}\ 17)} = g^{kj}g_{ik} = \delta_i^{\ j}.$$ □

Remark 3.4. Note that by equation (17), $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = g_{ij}$ and thus $\boldsymbol{e}^i \cdot \boldsymbol{e}^j = g^{ij}$.

Now, substituting the first of equations (29) and the second of equations (34) into equation (25), we find:

$$\vec{A} = \sum_i A_{(i)}\hat{e}_{(i)} = \sum_i \cancel{h_{(i)}}A^i\frac{1}{\cancel{h_{(i)}}}\boldsymbol{e}_i = A^i\boldsymbol{e}_i, \tag{36}$$

where the sum on i is now implied. Similarly, by substituting the first of equations (30) and the second of equations (35) into equation (25), we find:

$$\vec{A} = \sum_i A_{(i)}\hat{e}_{(i)} = \sum_i \cancel{h_{(i)}} g^{ij} A_j \frac{g_{ik}}{\cancel{h_{(i)}}} \boldsymbol{e}^k = \underbrace{g^{ij} g_{ik}}_{\delta^j{}_k} A_j \boldsymbol{e}^k = A_j \boldsymbol{e}^j. \tag{37}$$

Thus, we can find the i^{th} covariant component of a vector by calculating:

$$\vec{A} \cdot \boldsymbol{e}_i = A_j \underbrace{\boldsymbol{e}^j \cdot \boldsymbol{e}_i}_{\delta^j{}_i} = A_i. \tag{38}$$

using Theorem 3.2. Because we have used the covariant and contravariant basis vectors instead of the physical unit vectors, equation (38) is true regardless of whether the basis is orthogonal, and thus appears to give us a simpler prescription for finding vector components in a general basis than the algorithm outlined in equation (26). Alas, nothing is for free; computing the covariant and contravariant basis vectors from the physical unit vectors can consist of similar operations as row-reducing a matrix.

Finally, let us re-establish contact with the introductory remarks, and remind the reader that equation (2) is the velocity vector in its contravariant form whereas equation (3) is in its physical form, reproduced here for reference:

$$\begin{aligned} \vec{v}_{\text{con}} &= (\dot{r}, \dot{\vartheta}, \dot{\varphi}); \\ \vec{v}_{\text{phys}} &= (\dot{r}, r\dot{\vartheta}, r\sin\vartheta\,\dot{\varphi}). \end{aligned}$$

Given equation (31), the covarient form of the velocity vector in spherical polar coordinates is evidently:

$$\vec{v}_{\text{cov}} = (\dot{r}, r^2\dot{\vartheta}, r^2\sin^2\vartheta\,\dot{\varphi}). \tag{39}$$

3.2 The scalar and inner products

Definition 3.6. The *covariant* and *contravariant scalar products* of two rank 1 tensors, A and B, are defined as $g^{ij}A_iB_j$ and $g_{ij}A^iB^j$ respectively.

The covariant and contravariant scalar products actually have the same value, as seen by:

$$g^{ij}A_iB_j = A^jB_j \qquad \text{and} \qquad g_{ij}A^iB^j = A^ig_{ij}B^j = A^iB_i,$$

using equation (23). Similarly, one could show that the common value is A_iB^i. Therefore, the covariant and contravariant scalar product are referred to as simply the *scalar product*.

To find the scalar product in physical components, we start with equations (29) and (30) to write:

$$A^iB_i = \sum_i \frac{A_{(i)}}{h_{(i)}} \sum_j \frac{g_{ij}B_{(j)}}{h_{(j)}} = \sum_{ij} A_{(i)}B_{(j)} \underbrace{\frac{g_{ij}}{h_{(i)}h_{(j)}}}_{\hat{e}_{(i)}\cdot\hat{e}_{(j)}} = \Big(\sum_i A_{(i)}\hat{e}_{(i)}\Big)\cdot\Big(\sum_j B_{(j)}\hat{e}_{(j)}\Big)$$

$$= \vec{A}\cdot\vec{B} = \sum_i A_i B_i \quad \text{(last equality for orthogonal coordinates only)}, \tag{40}$$

using equations (17) and (25). Thus, the scalar product of two rank 1 tensors is just the ordinary dot product between two physical vectors in vector algebra.

Remark 3.5. The scalar product of two rank 1 tensors is really the contraction of the dyadic A_iB^j and thus, from Theorem 2.4, the scalar product is invariant under coordinate transformations.

Note that while both A_iB^j and A_iB_j are rank 2 tensors (the first mixed, the second covariant), only A_iB^i is invariant. To see that $\sum_i A_iB_i$ is *not* invariant, write:

$$\sum_k \tilde{A}_k\tilde{B}_k = \sum_k \frac{\partial x^i}{\partial \tilde{x}^k}A_i \frac{\partial x^j}{\partial \tilde{x}^k}B_j = \sum_k \frac{\partial x^i}{\partial \tilde{x}^k}\frac{\partial x^j}{\partial \tilde{x}^k}A_iB_j \neq \sum_i A_iB_i.$$

Unlike the proof to Theorem 2.4, $\partial x^i/\partial\tilde{x}^k$ and $\partial x^j/\partial\tilde{x}^k$ are not each other's inverse and there is no Kronecker delta δ^{ij} to be extracted, whence the inequality. Note that the middle two terms are actually triple sums, including the implicit sums on each of i and j.

Definition 3.7. The *covariant and contravariant scalar products* of two rank 2 tensors, S and T, are defined as $g^{ik}g^{jl}S_{kl}T_{ij}$ and $g_{ik}g_{jl}S^{ij}T^{kl}$ respectively.

Similar to the scalar product of two rank 1 tensors, these operations result in a scalar:

$$g^{ik}g^{jl}S_{kl}T_{ij} = S^{ij}T_{ij} = S_{kl}g^{ik}g^{jl}T_{ij} = S_{kl}T^{kl} = g_{ik}g_{jl}S^{ij}T^{kl}.$$

Thus, the covariant and contravariant scalar products of two rank 2 tensors give the same value and are collectively referred to as simply as the *scalar product*.

In terms of the physical components, use equation 32 to write:

$$S^{ij}T_{ij} = \sum_{ij}\frac{S_{(ij)}}{h_{(i)}h_{(j)}}\sum_{kl}\frac{g_{ik}g_{jl}T_{(kl)}}{h_{(k)}h_{(l)}} = \sum_{ijkl}S_{(ij)}T_{(kl)}\underbrace{\frac{g_{ik}}{h_{(i)}h_{(k)}}}_{\hat{e}_{(i)}\cdot\hat{e}_{(k)}}\underbrace{\frac{g_{jl}}{h_{(j)}h_{(l)}}}_{\hat{e}_{(j)}\cdot\hat{e}_{(l)}} \tag{41}$$

$$\equiv \mathsf{S}:\mathsf{T} = \sum_{ij}S_{(ij)}T_{(ij)} \quad \text{(last equality for orthogonal coordinates only)},$$

Here, I use the "colon product" notation frequently used in vector algebra. If S and T are matrices relative to an orthogonal basis, the colon product is simply the sum of the products of the $(i,j)^{\text{th}}$ element of S with the $(i,j)^{\text{th}}$ element of T, all "cross-terms" being zero. Note that if S (T) is the dyadic product of rank 1 tensors A and B (C and D), and thus $S^{ij} = A^iB^j$ ($T_{ij} = C_iD_j$), then we can rewrite equation (41) as:

$$S^{ij}T_{ij} = A^iB^jC_iD_j = (A^iC_i)(B^jD_j) = (\vec{A}\cdot\vec{C})(\vec{B}\cdot\vec{D}) = \mathsf{S}:\mathsf{T},$$

and, in "vector-speak", the *colon product* is sometimes referred to as the *double dot product*.

Now, scalar products are operations on two tensors of the same rank that yield a scalar. Similar operations on tensors of unequal rank yield a tensor of non-zero rank, the simplest example being the contraction of a rank 2 tensor, T, with a rank 1 tensor, A (or vice versa). In tensor notation, there are *sixteen* ways such a contraction can be represented: $T^{ij}A_j$, $T^i_{\ j}A^j$, $T_i^{\ j}A_j$, $T_{ij}A^j$, A_iT^{ij}, $A^iT_i^{\ j}$, $A_iT^i_{\ j}$, A^iT_{ij} plus eight more with the factors reversed (*e.g.*, $T^{ij}A_j = A_jT^{ij}$). Fortunately, these can be arranged naturally in two groups:

$$T^{ij}A_j \;=\; T^i_{\ j}A^j \;=\; g^{ik}T_k^{\ j}A_j \;=\; g^{ik}T_{kj}A^j \;\equiv\; (\mathsf{T}\cdot\vec{A})^i; \tag{42}$$

$$A_iT^{ij} \;=\; A^iT_i^{\ j} \;=\; g^{jk}A_iT^i_{\ k} \;=\; g^{jk}A^iT_{ik} \;\equiv\; (\vec{A}\cdot\mathsf{T})^j. \tag{43}$$

where, by example, we have defined the contravariant *inner product* between tensors of rank 1 and 2.

Definition 3.8. The *inner product* between two tensors of any rank is the contraction of the *inner* indices, namely the last index of the first tensor and the first index of the last tensor.

Thus, to know how to write $A^jT^i_{\ j}$ as an inner product, one first notices that, as written, it is the last index of the last tensor (T) that is involved in the contraction, not the first index. By commuting the tensors to get $T^i_{\ j}A^j$ (which doesn't change its value), the last index of the first tensor is now contracted with the first index of the last tensor and, with the tensors now in their "proper" order, we write down $T^i_{\ j}A^j = (\mathsf{T}\cdot\vec{A})^i$. Note that $(\vec{A}\cdot\mathsf{T})^i = A_jT^{ji} \neq A_jT^{ij}$ unless T is symmetric. Thus, the inner product does not generally commute: $\vec{A}\cdot\mathsf{T} \neq \mathsf{T}\cdot\vec{A}$.

In vector/matrix notation, $\mathsf{T}\cdot\vec{A}$ is the *right dot product* of the matrix T with the vector $\vec{A}$, and is equivalent to the matrix multiplication of the $m\times m$ matrix T on the left with the $1\times m$ *column* vector $\vec{A}$ on the right, yielding another $1\times m$ column vector, call it $\vec{B}$. In "bra-ket" notation, this is represented as $\mathsf{T}|A\rangle = |B\rangle$. Conversely, the *left dot product*, $\vec{A}\cdot\mathsf{T}$, is the matrix multiplication of the $m\times 1$ *row* vector on the left with the $m\times m$ matrix on the right, yielding another $m\times 1$ row vector. In "bra-ket" notation, this is represented as $\langle A|\mathsf{T} = \langle B|$.

The inner products defined in equations (42) and (43) are rank 1 contravariant tensors. In terms of the physical components, we have from equations (29) and (32):

$$\begin{aligned} T^{ij}A_j \;&=\; T^i_{\ j}A^j \;=\; (\mathsf{T}\cdot\vec{A})^i \;=\; \frac{1}{h_{(i)}}(\mathsf{T}\cdot\vec{A})_{(i)} \;=\; g_{jk}T^{ik}A^j \\ &=\; \sum_{jk} g_{jk}\frac{T_{(ik)}}{h_{(i)}h_{(k)}}\frac{A_{(j)}}{h_{(j)}} \;=\; \frac{1}{h_{(i)}}\sum_{jk}T_{(ik)}A_{(j)}\frac{g_{jk}}{h_{(j)}h_{(k)}} \;=\; \frac{1}{h_{(i)}}\sum_{jk}T_{(ik)}A_{(j)}\hat{e}_{(j)}\cdot\hat{e}_{(k)} \end{aligned} \tag{44}$$

$$\Rightarrow\quad (\mathsf{T}\cdot\vec{A})_{(i)} \;=\; \sum_{jk}T_{(ik)}A_{(j)}\hat{e}_{(j)}\cdot\hat{e}_{(k)}\;\left(=\;\sum_j T_{(ij)}A_{(j)}\right), \tag{45}$$

$$\begin{aligned} A_jT^{ji} \;&=\; A^jT_j^{\ i} \;=\; (\vec{A}\cdot\mathsf{T})^i \;=\; \frac{1}{h_{(i)}}(\vec{A}\cdot\mathsf{T})_{(i)} \;=\; A^jg_{jk}T^{ki} \\ &=\; \sum_{jk} g_{jk}\frac{A_{(j)}}{h_{(j)}}\frac{T_{(ki)}}{h_{(k)}h_{(i)}} \;=\; \frac{1}{h_{(i)}}\sum_{jk}A_{(j)}T_{(ki)}\frac{g_{jk}}{h_{(j)}h_{(k)}} \;=\; \frac{1}{h_{(i)}}\sum_{jk}A_{(j)}T_{(ki)}\hat{e}_{(j)}\cdot\hat{e}_{(k)} \end{aligned} \tag{46}$$

$$\Rightarrow \quad (\vec{A}\cdot\mathsf{T})_{(i)} \;=\; \sum_{jk} T_{(ki)}A_{(j)}\hat{e}_{(j)}\cdot\hat{e}_{(k)} \;\Bigg(=\; \sum_{j} T_{(ji)}A_{(j)}\Bigg), \tag{47}$$

where the equalities in parentheses are true for orthogonal coordinates only. Note that the only difference between equations (45) and (47) is the order of the indices on T.

3.3 Invariance of tensor expressions

The most important property of tensors is their ability to render an equation invariant under coordinate transformations. As indicated after equation (12), each term in a valid tensor expression must have the same free indices in the same positions. Thus, for example, $U_{ij}{}^{k} = V_i$ is invalid since each term does not have the same number of indices, though this equation could be rendered valid by contracting on two of the indices in U: $U_{ij}{}^{j} = V_i$. As a further example,

$$T_i{}^j{}_{kl} \;=\; A_i{}^j B_{kl} + C_m{}^{mj} D_{ikl}, \tag{48}$$

is a valid tensor expression, whereas,

$$S_i{}^{jk}{}_l \;=\; A_i{}^j B_{kl} + C_m{}^{mj} D_{ikl}, \tag{49}$$

is not valid because k is contravariant in S but covariant in the two terms on the RHS. Given the role of metrics in raising and lowering indices, we could "rescue" equation (49) by renaming the index k in S to n, say, and then multiplying the LHS by g_{kn}. Thus,

$$g_{kn}S_i{}^{jn}{}_l \;=\; A_i{}^j B_{kl} + C_m{}^{mj} D_{ikl}, \tag{50}$$

is now a valid tensor expression. And so it goes.

An immediate consequence of the rules for assembling a valid tensor expression is that it must have the same form in every coordinate system. Thus and for example, in transforming equation (48) from coordinate system x^i to coordinate system $\tilde{x}^{i'}$, we would write:

$$\begin{aligned}
\frac{\partial\tilde{x}^{i'}}{\partial x^i}\frac{\partial x^j}{\partial\tilde{x}^{j'}}\frac{\partial\tilde{x}^{k'}}{\partial x^k}\frac{\partial\tilde{x}^{l'}}{\partial x^l}\tilde{T}_{i'}{}^{j'}{}_{k'l'} \;&=\; \frac{\partial\tilde{x}^{i'}}{\partial x^i}\frac{\partial x^j}{\partial\tilde{x}^{j'}}A_{i'}{}^{j'}\frac{\partial\tilde{x}^{k'}}{\partial x^k}\frac{\partial\tilde{x}^{l'}}{\partial x^l}B_{k'l'} \\
&\quad + \underbrace{\frac{\partial\tilde{x}^{m'}}{\partial x^m}\frac{\partial x^m}{\partial\tilde{x}^{m'}}}_{1}\frac{\partial x^j}{\partial\tilde{x}^{j'}}C_{m'}{}^{m'j'}\frac{\partial\tilde{x}^{i'}}{\partial x^i}\frac{\partial\tilde{x}^{k'}}{\partial x^k}\frac{\partial\tilde{x}^{l'}}{\partial x^l}D_{i'k'l'} \\
\Rightarrow \quad \frac{\partial\tilde{x}^{i'}}{\partial x^i}\frac{\partial x^j}{\partial\tilde{x}^{j'}}\frac{\partial\tilde{x}^{k'}}{\partial x^k}\frac{\partial\tilde{x}^{l'}}{\partial x^l}\Big(\tilde{T}_{i'}{}^{j'}{}_{k'l'} - A_{i'}{}^{j'}B_{k'l'} - C_{m'}{}^{m'j'}D_{i'k'l'}\Big) \;&=\; 0.
\end{aligned}$$

Since no assumptions were made of the coordinate transformation factors (the derivatives) in front, this equation must be true for all possible factors, and thus can be true only if the quantity in parentheses is zero. Thus,

$$\tilde{T}_{i'}{}^{j'}{}_{k'l'} \;=\; A_{i'}{}^{j'}B_{k'l'} + C_{m'}{}^{m'j'}D_{i'k'l'}. \tag{51}$$

The fact that equation (51) has the identical form as equation (48) is what is meant by a tensor expression being invariant under coordinate transformations. Note that equation (49) would not transform in an invariant fashion, since the LHS would have different coordinate transformation factors than the RHS. Note further that the invariance of a tensor expression like equation (48) doesn't mean that each term remains unchanged under the coordinate transformation. Indeed, the components of most tensors will change under coordinate transformations. What doesn't change in a valid tensor expression is how the tensors are related to each other, with the changes to each tensor "cancelling out" from each term.

Finally, courses in tensor analysis often include some mention of the *quotient rule*, which has nothing to do with the quotient rule of single-variable calculus. Instead, it is an inverted restatement, of sorts, of what it means to be an invariant tensor expression for particularly simple expressions.

Theorem 3.3. (Quotient Rule) *If* A *and* B *are tensors, and if the expression* $\mathsf{A} = \mathsf{BT}$ *is invariant under coordinate transformation, then* T *is a tensor.*

Proof. Here we look at the special case:

$$A_i = B_j T^j{}_i. \tag{52}$$

The proof for tensors of general rank is more cumbersome and no more enlightening. Since equation (52) is invariant under coordinate transformations, we can write:

$$\tilde{B}_l \tilde{T}^l{}_k = \tilde{A}_k = \frac{\partial x^i}{\partial \tilde{x}^k} A_i = \frac{\partial x^i}{\partial \tilde{x}^k} B_j T^j{}_i = \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^l}{\partial x^j} \tilde{B}_l T^j{}_i$$

$$\Rightarrow \quad \tilde{B}_l \left(\tilde{T}^l{}_k - \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^l}{\partial x^j} T^j{}_i \right) = 0,$$

which must be true $\forall\ \tilde{B}_l$. This is possible only if the contents of the parentheses is zero, whence:

$$\tilde{T}^l{}_k = \frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial \tilde{x}^l}{\partial x^j} T^j{}_i,$$

and $T^j{}_i$ transforms as a rank 2 mixed tensor. □

3.4 The permutation tensors

Definition 3.9. The *Levi-Civita symbol*, ε^{ijk}, also known as the *three-dimensional permutation parameter*, is given by:

$$\varepsilon_{ijk} = \varepsilon^{ijk} = \begin{cases} 1 & \text{for } i, j, k \text{ an even permutation of } 1, 2, 3; \\ -1 & \text{for } i, j, k \text{ an odd permutation of } 1, 2, 3; \\ 0 & \text{if any of } i, j, k \text{ are the same.} \end{cases} \tag{53}$$

As it turns out, ε^{ijk} is *not* a tensor[7], though its indices will still participate in Einstein summation conventions where applicable and unless otherwise noted. As written in equation (53), there are two "flavours" of the Levi-Civita symbol—one with covariant-like indices, one with contravariant-like indices—which are used as convenient. Numerically, the two are equal.

[7]Technically, ε^{ijk} is a *pseudo*tensor, a distinction we will not need to make in this primer.

There are two very common uses for ε^{ijk}. First, it is used to represent vector cross-products:

$$C_k = (\vec{A} \times \vec{B})_k = \varepsilon_{ijk} A^i B^j. \tag{54}$$

In a similar vein, we shall see in §5.5 how it, or at least the closely related *permutation tensor* defined below, is used in the definition of the tensor curl.

Second, and most importantly, ε_{ijk} is used to represent determinants. If A is a 3×3 matrix, then its determinant, A, is given by:

$$A = \varepsilon^{ijk} A_{1i} A_{2j} A_{3k}, \tag{55}$$

which can be verified by direct application of equation (53). Indeed, determinants of higher dimensioned matrices may be represented by permutation parameters of higher rank and dimension. Thus, the determinant of an $m \times m$ matrix, B, is given by:

$$B = \varepsilon^{i_1 i_2 \dots i_m} B_{1i_1} B_{2i_2} \dots B_{mi_m},$$

where both the rank and dimension of $\varepsilon^{i_1 i_2 \dots i_m}$ is m (though the rank of matrix B is still 2).

Theorem 3.4. *Consider two 3-dimensional coordinate systems*[8], x^i *and* $\tilde{x}^{i'}$, *and let* $\mathcal{J}_{x,\tilde{x}}$ *be the* Jacobian determinant, *namely:*

$$\mathcal{J}_{x,\tilde{x}} = \left| \frac{\partial x^i}{\partial \tilde{x}^{i'}} \right| = \begin{vmatrix} \frac{\partial x^1}{\partial \tilde{x}^1} & \frac{\partial x^1}{\partial \tilde{x}^2} & \frac{\partial x^1}{\partial \tilde{x}^3} \\ \frac{\partial x^2}{\partial \tilde{x}^1} & \frac{\partial x^2}{\partial \tilde{x}^2} & \frac{\partial x^2}{\partial \tilde{x}^3} \\ \frac{\partial x^3}{\partial \tilde{x}^1} & \frac{\partial x^3}{\partial \tilde{x}^2} & \frac{\partial x^3}{\partial \tilde{x}^3} \end{vmatrix}$$

If A *and* $\tilde{\mathsf{A}}$ *are the same rank 2, 3-dimensional tensor in each of the two coordinate systems, and if* A *and* $\tilde{A}$ *are their respective determinants, then*

$$\mathcal{J}_{x,\tilde{x}} = \sqrt{\tilde{A}/A}. \tag{56}$$

Proof. We start with the observation that:

$$\varepsilon_{pqr} A = \varepsilon^{ijk} A_{pi} A_{qj} A_{rk}, \tag{57}$$

is logically equivalent to equation (55). This can be verified by direct substitution of all possible cases. Thus, if (p, q, r) is an even permutation of $(1, 2, 3)$, $\varepsilon_{pqr} = 1$ on the LHS while the matrix A on the RHS effectively undergoes an even number of row swaps, leaving the determinant unchanged. If (p, q, r) is an odd permutation of $(1, 2, 3)$, $\varepsilon_{pqr} = -1$ while A undergoes an odd number of row swaps, negating the determinant. In both cases, equation (57) is equivalent to equation (55). Finally, if any two of (p, q, r) are equal, $\varepsilon_{pqr} = 0$ and the determinant, now of a matrix with two identical rows, would also be zero.

Rewriting equation (57) in the $\tilde{x}$ coordinate system, we get:

$$\varepsilon_{p'q'r'} \tilde{A} = \varepsilon^{i'j'k'} \tilde{A}_{p'i'} \tilde{A}_{q'j'} \tilde{A}_{r'k'}$$

[8] The extension to m dimensions is straight-forward.

$$= \varepsilon^{i'j'k'} \frac{\partial x^p}{\partial \tilde{x}^{p'}} \frac{\partial x^i}{\partial \tilde{x}^{i'}} A_{pi} \frac{\partial x^q}{\partial \tilde{x}^{q'}} \frac{\partial x^j}{\partial \tilde{x}^{j'}} A_{qj} \frac{\partial x^r}{\partial \tilde{x}^{r'}} \frac{\partial x^k}{\partial \tilde{x}^{k'}} A_{rk}$$

$$= \underbrace{\varepsilon^{i'j'k'} \frac{\partial x^i}{\partial \tilde{x}^{i'}} \frac{\partial x^j}{\partial \tilde{x}^{j'}} \frac{\partial x^k}{\partial \tilde{x}^{k'}}}_{\varepsilon^{ijk}\mathcal{J}_{x,\tilde{x}}} A_{pi} A_{qj} A_{rk} \frac{\partial x^p}{\partial \tilde{x}^{p'}} \frac{\partial x^q}{\partial \tilde{x}^{q'}} \frac{\partial x^r}{\partial \tilde{x}^{r'}}$$

where the underbrace is a direct application of equation (57). Continuing...

$$\Rightarrow \quad \varepsilon_{p'q'r'} \tilde{A} = \underbrace{\varepsilon^{ijk} A_{pi} A_{qj} A_{rk}}_{\varepsilon_{pqr} A} \frac{\partial x^p}{\partial \tilde{x}^{p'}} \frac{\partial x^q}{\partial \tilde{x}^{q'}} \frac{\partial x^r}{\partial \tilde{x}^{r'}} \mathcal{J}_{x,\tilde{x}}$$

$$= \underbrace{\varepsilon_{pqr} \frac{\partial x^p}{\partial \tilde{x}^{p'}} \frac{\partial x^q}{\partial \tilde{x}^{q'}} \frac{\partial x^r}{\partial \tilde{x}^{r'}}}_{\varepsilon_{p'q'r'}\mathcal{J}_{x,\tilde{x}}} \mathcal{J}_{x,\tilde{x}}\, A$$

$$= \varepsilon_{p'q'r'} (\mathcal{J}_{x,\tilde{x}})^2 A$$

$$\Rightarrow \quad \varepsilon_{p'q'r'} \big(\tilde{A} - (\mathcal{J}_{x,\tilde{x}})^2 A\big) = 0 \quad \Rightarrow \quad \mathcal{J}_{x,\tilde{x}} = \sqrt{\tilde{A}/A}.$$

□

Theorem 3.5. *If $g = \det g_{ij}$ is the determinant of the metric tensor, then the entities:*

$$\epsilon^{ijk} = \frac{1}{\sqrt{g}} \varepsilon^{ijk}; \qquad \epsilon_{ijk} = \sqrt{g}\, \varepsilon_{ijk}, \tag{58}$$

are rank 3 tensors. ϵ^{ijk} and ϵ_{ijk} are known, respectively, as the contravariant and covariant permutation tensors.

Proof. Consider $\tilde{\epsilon}^{i'j'k'}$ in the $\tilde{x}^{i'}$ coordinate system. In transforming it to the x^i coordinate system, we would write:

$$\tilde{\epsilon}^{i'j'k'} \frac{\partial x^i}{\partial \tilde{x}^{i'}} \frac{\partial x^j}{\partial \tilde{x}^{j'}} \frac{\partial x^k}{\partial \tilde{x}^{k'}} = \frac{1}{\sqrt{\tilde{g}}} \underbrace{\varepsilon^{i'j'k'} \frac{\partial x^i}{\partial \tilde{x}^{i'}} \frac{\partial x^j}{\partial \tilde{x}^{j'}} \frac{\partial x^k}{\partial \tilde{x}^{k'}}}_{\varepsilon^{ijk}\mathcal{J}_{x,\tilde{x}}} = \frac{1}{\sqrt{\tilde{g}}} \varepsilon^{ijk} \sqrt{\frac{\tilde{g}}{g}} = \frac{1}{\sqrt{g}} \varepsilon^{ijk} = \epsilon^{ijk},$$

using first equation (57), then equation (56) with $A_{ij} = g_{ij}$. Thus, ϵ^{ijk} transforms like a rank 3 contravariant tensor. Further, its covariant conjugate is given by:

$$\epsilon_{pqr} = g_{pi} g_{qj} g_{rk} \epsilon^{ijk} = \frac{1}{\sqrt{g}} \underbrace{\varepsilon^{ijk} g_{pi} g_{qj} g_{rk}}_{\varepsilon_{pqr} g} = \sqrt{g} \varepsilon_{pqr},$$

and $\sqrt{g}\varepsilon_{pqr}$ is a rank 3 covariant tensor. □

4 Tensor derivatives

While the partial derivative of a scalar, $\partial f/\partial x^i$, is the prototypical covariant rank 1 tensor (equation 8), we get into trouble as soon as we try taking the derivative of a tensor of any higher rank. Consider the transformation of $\partial A^i/\partial x^j$ from the x^i coordinate system to $\tilde{x}^p$:

$$\frac{\partial \tilde{A}^p}{\partial \tilde{x}^q} = \frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial}{\partial x^j}\left(\frac{\partial \tilde{x}^p}{\partial x^i}A^i\right) = \frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial \tilde{x}^p}{\partial x^i}\frac{\partial A^i}{\partial x^j} + \frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial^2 \tilde{x}^p}{\partial x^j \partial x^i}A^i. \tag{59}$$

Now, if $\partial A^i/\partial x^j$ transformed as a tensor, we would have expected only the first term on the RHS. The presence of the second term means that $\partial A^i/\partial x^j$ is not a tensor, and we therefore need to generalise our definition of tensor differentiation if we want equations involving tensor derivatives to maintain their tensor invariance.

Before proceeding, however, let us consider an alternate and, as it turns out, incorrect approach. One might be tempted to write, as was I in preparing this primer:

$$\frac{\partial \tilde{A}^p}{\partial \tilde{x}^q} = \frac{\partial}{\partial \tilde{x}^q}\left(\frac{\partial \tilde{x}^p}{\partial x^i}A^i\right) = \frac{\partial \tilde{x}^p}{\partial x^i}\frac{\partial A^i}{\partial \tilde{x}^q} + A^i\frac{\partial}{\partial \tilde{x}^q}\frac{\partial \tilde{x}^p}{\partial x^i} = \frac{\partial \tilde{x}^p}{\partial x^i}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial A^i}{\partial x^j} + A^i\frac{\partial}{\partial x^i}\frac{\partial \tilde{x}^p}{\partial \tilde{x}^q}.$$

As written, the last term would be zero since $\partial \tilde{x}^p/\partial \tilde{x}^q = \delta^p{}_q$, and $\partial \delta^p{}_q/\partial x^i = 0$. This clearly disagrees with the last term in equation (59), so what gives? Since $\tilde{x}^q$ and x^i are *not* linearly independent of each other, their order of differentiation may *not* be swapped, and the second term on the RHS is bogus.

Returning to equation (59), rather than taking the derivative of a single vector component, take instead the derivative of the full vector $\vec{A} = A^i\boldsymbol{e}_i$ (equation 36):

$$\frac{\partial \vec{A}}{\partial x^j} = \frac{\partial (A^i\boldsymbol{e}_i)}{\partial x^j} = \frac{\partial A^i}{\partial x^j}\boldsymbol{e}_i + \frac{\partial \boldsymbol{e}_i}{\partial x^j}A^i. \tag{60}$$

The first term accounts for the rate of change of the vector component, A^i, from point to point, while the second term accounts for the rate of change of the basis vector, $\boldsymbol{e}_i$. Both terms are of equal importance and thus, to make progress in differentiating a general vector, we need to understand how to differentiate a basis vector.

4.1 "Christ-awful symbols"[9]

If $\boldsymbol{e}_i$ is one of m covariant basis vectors spanning an m-dimensional space, then $\partial \boldsymbol{e}_i/\partial x^j$ is a vector within that same m-dimensional space and therefore can be expressed as a linear combination of the m basis vectors:

Definition 4.1. The *Christoffel symbols of the second kind*, Γ^k_{ij}, are the components of the vector $\partial \boldsymbol{e}_i/\partial x^j$ relative to the basis $\boldsymbol{e}_k$. Thus,

$$\frac{\partial \boldsymbol{e}_i}{\partial x^j} = \Gamma^k_{ij}\boldsymbol{e}_k \tag{61}$$

[9] A tip of the hat to Professor Daniel Finley, the University of New Mexico.

To get an expression for Γ^l_{ij} by itself, we take the dot product of equation (61) with $\boldsymbol{e}^l$ (*cf.*, equation 38) to get:

$$\frac{\partial \boldsymbol{e}_i}{\partial x^j} \cdot \boldsymbol{e}^l = \Gamma^k_{ij} \underbrace{\boldsymbol{e}_k \cdot \boldsymbol{e}^l}_{\delta_k^{\,l}} = \Gamma^l_{ij}. \tag{62}$$

Theorem 4.1. *The Christoffel symbol of the second kind is symmetric in its lower indices.*

Proof. Recalling definition 3.5, we write:

$$\Gamma^l_{ij} = \frac{\partial \boldsymbol{e}_i}{\partial x^j} \cdot \boldsymbol{e}^l = \frac{\partial}{\partial x^j}\frac{\partial \vec{r}_x}{\partial x^i} \cdot \boldsymbol{e}^l = \frac{\partial}{\partial x^i}\frac{\partial \vec{r}_x}{\partial x^j} \cdot \boldsymbol{e}^l = \frac{\partial \boldsymbol{e}_j}{\partial x^i} \cdot \boldsymbol{e}^l = \Gamma^l_{ji}. \qquad \square$$

Definition 4.2. The *Christoffel symbols of the first kind*, $\Gamma_{ij\,k}$, are given by:

$$\Gamma_{ij\,k} = g_{lk}\Gamma^l_{ij}; \qquad \Gamma^l_{ij} = g^{lk}\Gamma_{ij\,k}. \tag{63}$$

Remark 4.1. It is easy to show that $\Gamma_{ij\,k}$ is symmetric in its first two indices.

A note on notation. For Christoffel symbols of the first kind, most authors use $[ij, k]$ instead of $\Gamma_{ij\,k}$, and for the second kind many use $\left\{ {l \atop ij} \right\}$ instead of Γ^l_{ij}. While the Christoffel symbols are not tensors (as will be shown later) and thus up/down indices do not indicate contravariance/covariance, I prefer the Γ notation because, by definition, the two kinds of Christoffel symbols are related through the metric just like conjugate tensors. While it is only the non-symmetric index that can be raised or lowered on a Christoffel symbol, this still makes this notation useful. Further, we shall find it practical to allow Christoffel symbols to participate in the Einstein summation convention (as in equations 61, 62, and 63), and thus we do not enclose their indices in parentheses.

For the cognoscenti, it is acknowledged that the Γ notation is normally reserved for a quantity called the *affine connexion*, a concept from differential geometry and manifold theory. It plays an important role in General Relativity, where one can show that the affine connexion, Γ^l_{ij}, is equal to the Christoffel symbol of the second kind, $\left\{ {l \atop ij} \right\}$ (Weinberg, *Gravitation and Cosmology*, ISBN 0-471-92567-5, pp. 100–101), whence my inclination to borrow the Γ notation for Christoffel symbols.

Using equations (62) and (63), we can write:

$$\Gamma_{ij\,k} = g_{lk}\Gamma^l_{ij} = g_{lk}\boldsymbol{e}^l \cdot \frac{\partial \boldsymbol{e}_i}{\partial x^j} = \boldsymbol{e}_k \cdot \frac{\partial \boldsymbol{e}_i}{\partial x^j}. \tag{64}$$

Now from remark 3.4, we have $g_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{e}_j$, and thus,

$$\frac{\partial g_{ij}}{\partial x^k} = \boldsymbol{e}_i \cdot \frac{\partial \boldsymbol{e}_j}{\partial x^k} + \boldsymbol{e}_j \cdot \frac{\partial \boldsymbol{e}_i}{\partial x^k} = \Gamma_{jk\,i} + \Gamma_{ik\,j}, \tag{65}$$

using equation (64). Permuting the indices on equation (65) twice, we get:

$$\frac{\partial g_{ki}}{\partial x^j} = \Gamma_{ij\,k} + \Gamma_{kj\,i}; \qquad \text{and} \qquad \frac{\partial g_{jk}}{\partial x^i} = \Gamma_{ki\,j} + \Gamma_{ji\,k}. \tag{66}$$

(x^1, x^2, x^3)	$\Gamma_{22\,1}$	$\Gamma_{33\,1}$	$\Gamma_{12\,2}$	$\Gamma_{21\,2}$	$\Gamma_{33\,2}$	$\Gamma_{13\,3}$	$\Gamma_{23\,3}$	$\Gamma_{31\,3}$	$\Gamma_{32\,3}$
(z, ϱ, φ)	0	0	0	0	$-\varrho$	0	ϱ	0	ϱ
(r, ϑ, φ)	$-r$	$-r\sin^2\vartheta$	r	r	$-\dfrac{r^2}{2}\sin 2\vartheta$	$r\sin^2\vartheta$	$\dfrac{r^2}{2}\sin 2\vartheta$	$r\sin^2\vartheta$	$\dfrac{r^2}{2}\sin 2\vartheta$
	Γ^1_{22}	Γ^1_{33}	Γ^2_{12}	Γ^2_{21}	Γ^2_{33}	Γ^3_{13}	Γ^3_{23}	Γ^3_{31}	Γ^3_{32}
(z, ϱ, φ)	0	0	0	0	$-\varrho$	0	$\dfrac{1}{\varrho}$	0	$\dfrac{1}{\varrho}$
(r, ϑ, φ)	$-r$	$-r\sin^2\vartheta$	$\dfrac{1}{r}$	$\dfrac{1}{r}$	$-\frac{1}{2}\sin 2\vartheta$	$\dfrac{1}{r}$	$\cot\vartheta$	$\dfrac{1}{r}$	$\cot\vartheta$

Table 2: Christoffel symbols for cylindrical and spherical polar coordinates, as given by equation (68). All Christoffel symbols not listed are zero.

By combining equations (65) and (66), and exploiting the symmetry of the first two indices on the Christoffel symbols, one can easily show that:

$$\Gamma_{ij\,k} = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right) \quad \Rightarrow \quad \Gamma^l_{ij} = \frac{g^{lk}}{2}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right). \tag{67}$$

For a general 3-D coordinate system, there are 27 Christoffel symbols of each kind (15 of which are independent), each with as many as nine terms, whence the title of this subsection. However, for orthogonal coordinates where $g_{ij} \propto \delta_{ij}$ and $g_{ii} = h^2_{(i)} = 1/g^{ii}$, things get much simpler. In this case, Christoffel symbols fall into three categories as follows (no sum convention):

$$\left.\begin{aligned}
\Gamma_{ij\,i} &= g_{ii}\Gamma^i_{ij} = \frac{1}{2}\frac{\partial g_{ii}}{\partial x^j} && \text{for } i,j = 1,2,3 && \text{(15 components)} \\
\Gamma_{ii\,j} &= g_{jj}\Gamma^j_{ii} = -\frac{1}{2}\frac{\partial g_{ii}}{\partial x^j} && \text{for } i \neq j && \text{(6 components)} \\
\Gamma_{ij\,k} &= g_{kk}\Gamma^k_{ij} = 0 && \text{for } i \neq j \neq k && \text{(6 components)}
\end{aligned}\right\} \tag{68}$$

In this primer, our examples are restricted to the most common orthogonal coordinate systems in 3-space, namely Cartesian, cylindrical, and spherical polar where the Christoffel symbols aren't so bad to deal with. For Cartesian coordinates, all Christoffel symbols are zero, while those for cylindrical and spherical polar coordinates are given in Table 2.

To determine how Christoffel symbols transform under coordinate transformations, we first consider how the metric derivative transforms from coordinate system x^i to $\tilde{x}^p$. Thus,

$$\frac{\partial \tilde{g}_{pq}}{\partial \tilde{x}^r} = \frac{\partial}{\partial \tilde{x}^r}\left(g_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\right) = \frac{\partial g_{ij}}{\partial x^k}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} + g_{ij}\frac{\partial^2 x^i}{\partial \tilde{x}^r \partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} + g_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial^2 x^j}{\partial \tilde{x}^r \partial \tilde{x}^q}. \tag{69}$$

Permute indices $(p, q, r) \to (q, r, p) \to (r, p, q)$ and $(i, j, k) \to (j, k, i) \to (k, i, j)$ to get:

$$\frac{\partial \tilde{g}_{qr}}{\partial \tilde{x}^p} = \frac{\partial g_{jk}}{\partial x^i}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{jk}\frac{\partial^2 x^j}{\partial \tilde{x}^p \partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{jk}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial^2 x^k}{\partial \tilde{x}^p \partial \tilde{x}^r}; \tag{70}$$

$$\frac{\partial \tilde{g}_{rp}}{\partial \tilde{x}^q} = \frac{\partial g_{ki}}{\partial x^j}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p} + g_{ki}\frac{\partial^2 x^k}{\partial \tilde{x}^q \partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p} + g_{ki}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial^2 x^i}{\partial \tilde{x}^q \partial \tilde{x}^p}, \tag{71}$$

Using equations (69)–(71), we construct $\Gamma_{pq\,r}$ according to the first of equations (67) to get:

$$\begin{aligned}\tilde{\Gamma}_{pq\,r} = \frac{1}{2}\bigg(&\frac{\partial g_{jk}}{\partial x^i}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{ij}\frac{\partial^2 x^i}{\partial \tilde{x}^p \partial \tilde{x}^q}\frac{\partial x^j}{\partial \tilde{x}^r} + g_{ij}\frac{\partial x^i}{\partial \tilde{x}^q}\frac{\partial^2 x^j}{\partial \tilde{x}^p \partial \tilde{x}^r} \\ &+ \frac{\partial g_{ki}}{\partial x^j}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p} + g_{ij}\frac{\partial^2 x^i}{\partial \tilde{x}^q \partial \tilde{x}^r}\frac{\partial x^j}{\partial \tilde{x}^p} + g_{ij}\frac{\partial x^i}{\partial \tilde{x}^r}\frac{\partial^2 x^j}{\partial \tilde{x}^q \partial \tilde{x}^p} \\ &- \frac{\partial g_{ij}}{\partial x^k}\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} - g_{ij}\frac{\partial^2 x^i}{\partial \tilde{x}^r \partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q} - g_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial^2 x^j}{\partial \tilde{x}^r \partial \tilde{x}^q}\bigg),\end{aligned} \tag{72}$$

where the indices in the terms proportional to metric derivatives have been left unaltered, and where the dummy indices in the terms proportional to the metric have been renamed i and j. Now, since the metric is symmetric, $g_{ij} = g_{ji}$, we can write the third term on the RHS of equation (72) as:

$$g_{ij}\frac{\partial x^i}{\partial \tilde{x}^q}\frac{\partial^2 x^j}{\partial \tilde{x}^p \partial \tilde{x}^r} = g_{ji}\frac{\partial x^i}{\partial \tilde{x}^q}\frac{\partial^2 x^j}{\partial \tilde{x}^p \partial \tilde{x}^r} = g_{ij}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial^2 x^i}{\partial \tilde{x}^p \partial \tilde{x}^r},$$

where the dummy indices were once again renamed after the second equal sign. Performing the same manipulations to the sixth and ninth terms on the RHS of equation (72), one finds that the third and eighth terms cancel, the fifth and ninth terms cancel, and the second and sixth terms are the same. Thus, equation (72) simplifies mercifully to:

$$\begin{aligned}\tilde{\Gamma}_{pq\,r} &= \frac{1}{2}\bigg(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\bigg)\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{ij}\frac{\partial x^j}{\partial \tilde{x}^r}\frac{\partial^2 x^i}{\partial \tilde{x}^p \partial \tilde{x}^q} \\ &= \Gamma_{ij\,k}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial x^k}{\partial \tilde{x}^r} + g_{ij}\frac{\partial x^j}{\partial \tilde{x}^r}\frac{\partial^2 x^i}{\partial \tilde{x}^p \partial \tilde{x}^q},\end{aligned} \tag{73}$$

and the Christoffel symbol of the first kind does not transform like a tensor because of the second term on the right hand side. Like the ordinary derivative of a first rank tensor (equation 59), what prevents $\Gamma_{ij\,k}$ from transforming like a tensor is a term proportional to the second derivative of the coordinates. This important coincidence will be exploited when we define the *covariant derivative* in the next subsection.

Finally, to determine how the Christoffel symbol of the second kind transforms, we need only multiply equation (73) by:

$$\tilde{g}^{rs} = g^{kl}\frac{\partial \tilde{x}^r}{\partial x^k}\frac{\partial \tilde{x}^s}{\partial x^l},$$

to get:

$$\tilde{\Gamma}^s_{pq} = \Gamma^l_{ij}\frac{\partial x^i}{\partial \tilde{x}^p}\frac{\partial x^j}{\partial \tilde{x}^q}\underbrace{\frac{\partial x^k}{\partial \tilde{x}^r}\frac{\partial \tilde{x}^r}{\partial x^k}}_{1}\frac{\partial \tilde{x}^s}{\partial x^l} + \underbrace{g^{kl}\,\underbrace{g_{ij}\underbrace{\frac{\partial x^j}{\partial \tilde{x}^r}\frac{\partial \tilde{x}^r}{\partial x^k}}_{\delta^j{}_k}}_{g_{ik}}}_{\delta^l_i}\frac{\partial \tilde{x}^s}{\partial x^l}\frac{\partial^2 x^i}{\partial \tilde{x}^p \partial \tilde{x}^q}$$

$$\Rightarrow \quad \tilde{\Gamma}^s_{pq} = \Gamma^l_{ij} \frac{\partial x^i}{\partial \tilde{x}^p} \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^s}{\partial x^l} + \frac{\partial \tilde{x}^s}{\partial x^l} \frac{\partial^2 x^l}{\partial \tilde{x}^p \partial \tilde{x}^q}. \tag{74}$$

Once again, all that stops Γ^l_{ij} from transforming like a tensor is the second term proportional to the second derivative.

4.2 Covariant derivative

Substituting equation (61) into equation (60), we get:

$$\frac{\partial \vec{A}}{\partial x^j} = \frac{\partial A^i}{\partial x^j} \boldsymbol{e}_i + \Gamma^k_{ij} \boldsymbol{e}_k A^i = \left(\frac{\partial A^i}{\partial x^j} + \Gamma^i_{jk} A^k \right) \boldsymbol{e}_i, \tag{75}$$

where the names of the dummy indices i and k are swapped after the second equality.

Definition 4.3. The *covarient derivative* of a contravariant vector, A^i, is given by:

$$\nabla_j A^i \equiv \frac{\partial A^i}{\partial x^j} + \Gamma^i_{jk} A^k, \tag{76}$$

where the adjective *covariant* refers to the fact that the index on the differentiation operator (j) is in the covariant (lower) position.

Thus, equation (75) becomes:

$$\frac{\partial \vec{A}}{\partial x^j} = \nabla_j A^i \boldsymbol{e}_i. \tag{77}$$

Thus, the i^{th} contravariant component of the vector $\partial \vec{A} / \partial x^j$ relative to the covariant basis $\boldsymbol{e}_i$ is the covariant derivative of the i^{th} contravariant component of the vector, A^i, with respect to the coordinate x^j. In general, covariant derivatives are much more cumbersome than partial derivatives as the covariant derivative of any one tensor component involves *all* tensor components for non-zero Christoffel symbols. Only for Cartesian coordinates—where all Christoffel symbols are zero—do covariant derivatives reduce to ordinary partial derivatives.

Consider now the transformation of the covariant derivative of a contravariant vector from the coordinate system x^i to $\tilde{x}^p$. Thus, using equations (59) and (74), we have:

$$\begin{aligned}
\tilde{\nabla}_q \tilde{A}^p &= \frac{\partial \tilde{A}^p}{\partial \tilde{x}^q} + \tilde{\Gamma}^p_{qr} \tilde{A}^r \\
&= \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^p}{\partial x^i} \frac{\partial A^i}{\partial x^j} + \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial^2 \tilde{x}^p}{\partial x^j \partial x^i} A^i + \left(\Gamma^i_{jk} \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial x^k}{\partial \tilde{x}^r} \frac{\partial \tilde{x}^p}{\partial x^i} + \frac{\partial \tilde{x}^p}{\partial x^i} \frac{\partial^2 x^i}{\partial \tilde{x}^q \partial \tilde{x}^r} \right) \frac{\partial \tilde{x}^r}{\partial x^l} A^l \\
&= \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^p}{\partial x^i} \frac{\partial A^i}{\partial x^j} + \Gamma^i_{jk} \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial \tilde{x}^p}{\partial x^i} \underbrace{\frac{\partial x^k}{\partial \tilde{x}^r} \frac{\partial \tilde{x}^r}{\partial x^l}}_{\delta^k_{\ l}} A^l + \frac{\partial x^j}{\partial \tilde{x}^q} \frac{\partial^2 \tilde{x}^p}{\partial x^j \partial x^i} A^i + \frac{\partial \tilde{x}^p}{\partial x^i} \frac{\partial^2 x^i}{\partial \tilde{x}^q \partial \tilde{x}^r} \frac{\partial \tilde{x}^r}{\partial x^l} A^l.
\end{aligned} \tag{78}$$

Now for some fun. Remembering we can swap the order of differentiation between linearly independent quantities, and exploiting our freedom to rename dummy indices at whim, we rewrite the last term of equation (78) as:

$$\frac{\partial \tilde{x}^p}{\partial x^i}\frac{\partial \tilde{x}^r}{\partial x^l}\frac{\partial}{\partial \tilde{x}^r}\frac{\partial x^i}{\partial \tilde{x}^q}A^l = \frac{\partial \tilde{x}^p}{\partial x^i}\frac{\partial}{\partial x^l}\frac{\partial x^i}{\partial \tilde{x}^q}A^l = \left(\frac{\partial}{\partial x^l}\left[\frac{\partial \tilde{x}^p}{\partial x^i}\frac{\partial x^i}{\partial \tilde{x}^q}\right] - \frac{\partial x^i}{\partial \tilde{x}^q}\frac{\partial}{\partial x^l}\frac{\partial \tilde{x}^p}{\partial x^i}\right)A^l$$

$$= \left(\cancelto{0}{\frac{\partial}{\partial x^l}\frac{\partial \tilde{x}^p}{\partial \tilde{x}^q}} - \frac{\partial x^i}{\partial \tilde{x}^q}\frac{\partial^2 \tilde{x}^p}{\partial x^l \partial x^i}\right)A^l = -\frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial^2 \tilde{x}^p}{\partial x^j \partial x^i}A^i,$$

where the term set to zero is zero because $\partial\tilde{x}^p/\partial\tilde{x}^q = \delta^p{}_q$ whose derivative is zero. Thus, the last two terms in equation (78) cancel, which then becomes:

$$\tilde{\nabla}_q\tilde{A}^p = \frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial \tilde{x}^p}{\partial x^i}\left(\frac{\partial A^i}{\partial x^j} + \Gamma^i_{jk}A^k\right) = \frac{\partial x^j}{\partial \tilde{x}^q}\frac{\partial \tilde{x}^p}{\partial x^i}\nabla_j A^i,$$

and the covariant derivative of a contravariant vector is a mixed rank 2 tensor.

Our results are for a contravariant vector because we chose to represent the vector $\vec{A} = A^i\boldsymbol{e}_i$ in equation (60). If, instead, we chose $\vec{A} = A_i\boldsymbol{e}^i$, we would have found that:

$$\frac{\partial \boldsymbol{e}^i}{\partial x^j} = -\Gamma^i_{jk}\boldsymbol{e}^k \tag{79}$$

instead of equation (61), and this would have lead to:

Definition 4.4. The *covarient derivative* of a covariant vector, A_i, is given by:

$$\nabla_j A_i \equiv \frac{\partial A_i}{\partial x^j} - \Gamma^k_{ij}A_k. \tag{80}$$

In the same way we proved $\nabla_j A^i$ is a rank 2 mixed tensor, one can show that $\nabla_j A_i$ is a rank 2 covariant tensor. Notice the minus sign in equation (80), which distinguishes covariant derivatives of covariant vectors from covariant derivatives of contravariant vectors. In principle, contravariant derivatives (∇^j) can also be defined for both covariant and contravariant vectors, though these are rarely used in practise.

A note on notation. While most branches of mathematics have managed to converge to a prevailing nomenclature for its primary constructs, tensor calculus is not one of them. To wit,

$$\nabla_k A_i = A_{i;k} = A_{i/k} = A_j|_k,$$

are all frequently-used notations for covariant derivatives of covariant vectors. Some authors even use $A_{i,k}$ (comma, not a semi-colon), which can be very confusing since other authors use the comma notation for partial derivatives:

$$\partial_k A_i = A_{i,k} = \frac{\partial A_i}{\partial x^k}.$$

In this primer, I shall use exclusively the nabla notation ($\nabla_k A_i$) for covariant derivatives, and either the full Leibniz notation for partial derivatives ($\partial A_i/\partial x^k$) as I have been doing so far, or the abbreviated Leibniz notation ($\partial_k A_i$) when convenient. In particular, I avoid like the plague the semi-colon ($A_{i;k}$) and comma ($A_{i,k}$) conventions.

Theorem 4.2. *The covariant derivatives of the contravariant and covariant basis vectors are zero.*

Proof. Starting with equation (76),

$$\nabla_j \boldsymbol{e}^i = \partial_j \boldsymbol{e}^i + \Gamma^i_{jk} \boldsymbol{e}^k = 0,$$

from equation (79). Similarly, starting with equation (80),

$$\nabla_j \boldsymbol{e}_i = \partial_j \boldsymbol{e}_i - \Gamma^k_{ij} \boldsymbol{e}_k = 0,$$

from equation (61). □

Covariant derivatives of higher rank tensors are often required, and are listed below without proof. Covariant derivatives of all three types of rank 2 tensors are:

$$\nabla_k T^{ij} = \partial_k T^{ij} + \Gamma^i_{k\alpha} T^{\alpha j} + \Gamma^j_{k\alpha} T^{i\alpha} \qquad (\mathsf{T} \text{ contravariant}); \tag{81}$$

$$\nabla_k T^i_{\ j} = \partial_k T^i_{\ j} + \Gamma^i_{k\alpha} T^\alpha_{\ j} - \Gamma^\alpha_{kj} T^i_{\ \alpha} \qquad (\mathsf{T} \text{ mixed}); \tag{82}$$

$$\nabla_k T_i^{\ j} = \partial_k T_i^{\ j} - \Gamma^\alpha_{ki} T_\alpha^{\ j} + \Gamma^j_{k\alpha} T_i^{\ \alpha} \qquad (\mathsf{T} \text{ mixed}); \tag{83}$$

$$\nabla_k T_{ij} = \partial_k T_{ij} - \Gamma^\alpha_{ki} T_{\alpha j} - \Gamma^\alpha_{kj} T_{i\alpha} \qquad (\mathsf{T} \text{ covariant}). \tag{84}$$

More generally, the covariant derivative of a mixed tensor of rank $p + q = n$ is given by:

$$\begin{aligned} \nabla_k T_{i_1 \dots i_p}{}^{j_1 \dots j_q} = \partial_k T_{i_1 \dots i_p}{}^{j_1 \dots j_q} &- \Gamma^\alpha_{ki_1} T_{\alpha, i_2 \dots i_p}{}^{j_1 \dots j_q} - \dots - \Gamma^\alpha_{ki_p} T_{i_1 \dots i_{p-1}\alpha}{}^{j_1 \dots j_q} \\ &+ \Gamma^{j_1}_{k\alpha} T_{i_1 \dots i_p}{}^{\alpha, j_2 \dots j_q} + \dots + \Gamma^{j_q}_{k\alpha} T_{i_1 \dots i_p}{}^{j_1 \dots j_{q-1}\alpha}. \end{aligned} \tag{85}$$

In general, a covariant derivative of a rank n tensor with p covariant indices and q contravariant indices will itself be a rank $n+1$ tensor with $p+1$ covariant indices and q contravariant indices.

Theorem 4.3. Summation rule: *If* A *and* B *are two tensors of the same rank, dimensionality, and type, then* $\nabla_i(A + B) = \nabla_i A + \nabla_i B$*, where the indices have been omitted for generality.*

Proof. The proof for a general tensor is awkward, and no more illuminating than for the special case of two rank 1 contravariant tensors:

$$\begin{aligned} \nabla_i (A^j + B^j) &= \partial_i (A^j + B^j) + \Gamma^j_{ik}(A^k + B^k) = (\partial_i A^j + \Gamma^j_{ik} A^k) + (\partial_i B^j + \Gamma^j_{ik} B^k) \\ &= \nabla_i A^j + \nabla_i B^j. \end{aligned}$$

□

Theorem 4.4. Product rule: *If* A *and* B *are two tensors of possibly different rank, dimensionality, and type, then* $\nabla_i(AB) = A\nabla_i B + B\nabla_i A$*.*

Proof. Consider a rank 2 contravariant tensor, A^{jk} and a rank 1 covariant tensor, B_l. The product $A^{jk}B_l$ is a mixed rank 3 tensor and its covariant derivative is given by an application of equation (85):

$$\begin{aligned}\nabla_i(A^{jk}B_l) &= \partial_i(A^{jk}B_l) + \Gamma^j_{i\alpha}A^{\alpha k}B_l + \Gamma^k_{i\alpha}A^{j\alpha}B_l - \Gamma^\alpha_{il}A^{jk}B_\alpha \\ &= A^{jk}\partial_i B_l + B_l\partial_i A^{jk} + B_l\Gamma^j_{i\alpha}A^{\alpha k} + B_l\Gamma^k_{i\alpha}A^{j\alpha} - A^{jk}\Gamma^\alpha_{il}B_\alpha \\ &= A^{jk}(\partial_i B_l - \Gamma^\alpha_{il}B_\alpha) + B_l(\partial_i A^{jk} + \Gamma^j_{i\alpha}A^{\alpha k} + \Gamma^k_{i\alpha}A^{j\alpha}) \\ &= A^{jk}\nabla_i B_l + B_l\nabla_i A^{jk}.\end{aligned}$$

The proof for tensors of general rank follows the same lines, though is much more cumbersome to write down. □

Theorem 4.5. Ricci's Theorem*: The covariant derivative of the metric and its inverse vanish.*

Proof. From equation (84), we have:

$$\begin{aligned}\nabla_k g_{ij} &= \partial_k g_{ij} - \Gamma^\alpha_{ki}g_{\alpha j} - \Gamma^\alpha_{kj}g_{i\alpha} = \partial_k g_{ij} - \Gamma_{ki\,j} - \Gamma_{kj\,i} \\ &= \partial_k g_{ij} - \frac{1}{2}(\partial_k g_{ij} + \partial_i g_{jk} - \partial_j g_{ki}) - \frac{1}{2}(\partial_k g_{ji} + \partial_j g_{ik} - \partial_i g_{kj}) = 0,\end{aligned}$$

owing to the symmetry of the metric tensor. As for its inverse, g^{ij}, we first note from equation (83) that:

$$\nabla_k \delta_i^{\ j} = \partial_k \delta_i^{\ j} - \Gamma^\alpha_{ki}\delta_\alpha^{\ j} + \Gamma^j_{k\alpha}\delta_i^{\ \alpha} = 0 - \Gamma^j_{ki} + \Gamma^j_{ki} = 0,$$

$$\begin{aligned}\Rightarrow \quad \nabla_k(g_{i\alpha}g^{\alpha j}) &= \nabla_k \delta_i^{\ j} = 0 \\ &= g_{i\alpha}\nabla_k g^{\alpha j} + g^{\alpha j}\cancelto{0}{\nabla_k g_{i\alpha}} = g_{i\alpha}\nabla_k g^{\alpha j},\end{aligned}$$

using the product rule (Theorem 4.4). We can't conclude directly from this that $\nabla_k g^{\alpha j} = 0$ since we are not allowed to "divide through" by a tensor component, particularly one implicated in a summation. We can, however, multiply through by the tensor's inverse to get:

$$\Rightarrow \quad \underbrace{g^{\beta i}g_{i\alpha}}_{\delta^\beta{}_\alpha}\nabla_k g^{\alpha j} = g^{\beta i}(0) = 0 \quad \Rightarrow \quad \nabla_k g^{\beta j} = 0.$$ □

This is not to say that the metric is a constant function of the coordinates. Indeed, except for Cartesian coordinates, $\partial_k g_{ij}$ will, in general, not be zero. The covariant derivative not only measures how functions change from point to point, it also takes into account how the basis vectors themselves change (in magnitude, direction, or both) as a function of position, and it is the sum of these two changes that is zero for the metric tensor.

Corollary 4.1. *The metric "passes through" a covariant derivative operator. That is:*

$$\nabla_k T^i_{\ j} = \nabla_k g_{\alpha j}T^{i\alpha} = g_{\alpha j}\nabla_k T^{i\alpha}.$$

Proof. From the product rule for covariant derivatives (Theorem 4.4),

$$\nabla_k g_{\alpha j} T^{i\alpha} = g_{j\alpha} \nabla_k T^{i\alpha} + T^{i\alpha} \cancelto{0}{\nabla_k g_{\alpha j}} = g_{\alpha j} \nabla_k T^{i\alpha},$$

as desired. The same can be shown for the metric inverse, g^{ij}. □

Now if one can take a covariant derivative once, second order and higher covariant derivatives must also be possible. Thus and for example, taking the covariant derivative of equation (76), we get:

$$\begin{aligned}
\nabla_k \nabla_j A^i \equiv \nabla_{jk} A^i &= \nabla_k(\partial_j A^i + \Gamma^i_{j\alpha} A^\alpha) \equiv \nabla_k B_j^{\ i} \\
&= \partial_k B_j^{\ i} - \Gamma^\beta_{kj} B_\beta^{\ i} + \Gamma^i_{k\beta} B_j^{\ \beta} \\
&= \partial_k \partial_j A^i + \partial_k(\Gamma^i_{j\alpha} A^\alpha) - \Gamma^\beta_{kj}(\partial_\beta A^i + \Gamma^i_{\beta\alpha} A^\alpha) + \Gamma^i_{k\beta}(\partial_j A^\beta + \Gamma^\beta_{j\alpha} A^\alpha) \\
&= \partial_{jk} A^i + \Gamma^i_{j\alpha} \partial_k A^\alpha - \Gamma^\alpha_{kj} \partial_\alpha A^i + \Gamma^i_{k\alpha} \partial_j A^\alpha + A^\alpha(\partial_k \Gamma^i_{j\alpha} - \Gamma^\beta_{kj} \Gamma^i_{\beta\alpha} + \Gamma^i_{k\beta} \Gamma^\beta_{j\alpha}),
\end{aligned} \tag{86}$$

not a pretty sight. Again, for Cartesian coordinates, all but the first term disappears. For other coordinate systems, it can get ugly fast. Four of the seven terms are single sums, two are double sums and there can be as many as 31 terms to calculate for every one of 27 components in 3-space! Aggravating the situation is the fact that the order of covariant differentiation cannot, in general, be swapped. Thus, while $\partial_{jk} A^i = \partial_{kj} A^i$ so long as x^j and x^k are independent, $\nabla_{jk} A^i \neq \nabla_{kj} A^i$ because the last term on the RHS of equation (86) is not, in general, symmetric in the interchange of j and k[10].

[10] Note, however, that all other terms are symmetric (second and fourth together), including the fifth term which can be shown to be symmetric in j and k with the aid of equation 62.

5 Connexion to vector calculus

For many applications, it is useful to apply the ideas of tensor analysis to three-dimensional vector analysis. For starters, this gives us another way to see how the metric factors arise in the definitions of the gradient of a scalar, and the divergence and curl of a vector. More importantly, it allows us to write down covariant expressions for more complicated derivatives, such as the gradient of a vector, and to prove certain tensor identities.

5.1 Gradient of a scalar

The prototypical covariant tensor of rank 1, $\partial_i f$ whose transformation properties are given by equation (8), was referred to as the "gradient", though it is not the physical gradient we would measure. The physical gradient, ∇f, is related to the covariant gradient, $\partial_i f$, by equation (31), namely $(\nabla f)_{(i)} = (\partial_i f)/h_{(i)}$ for orthogonal coordinates. Thus,

$$\nabla f \;=\; \left(\frac{1}{h_{(1)}}\partial_1 f,\; \frac{1}{h_{(2)}}\partial_2 f,\; \frac{1}{h_{(3)}}\partial_3 f\right). \tag{87}$$

For non-orthogonal coordinates, one would use equation (30) which would yield a somewhat more complicated expression for the gradient where each component becomes a sum.

5.2 Divergence of a vector

Consider the contraction of the covariant derivative of a contravariant vector,

$$\nabla_i A^i \;=\; \partial_i A^i + \Gamma^i_{ij} A^j. \tag{88}$$

Now, from the second of equations (67),

$$\begin{aligned}
\Gamma^i_{ij} &= \tfrac{1}{2}\, g^{ki}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \;=\; \tfrac{1}{2}(g^{ki}\partial_i g_{jk} + g^{ki}\partial_j g_{ki} - g^{ki}\partial_k g_{ij}) \\
&= \tfrac{1}{2}(g^{ki}\partial_i g_{jk} + g^{ki}\partial_j g_{ki} - g^{ik}\partial_i g_{kj}) \qquad \text{(swap dummy indices in 3}^{\text{rd}}\text{ term)} \\
&= \tfrac{1}{2}(g^{ki}\partial_i g_{jk} + g^{ki}\partial_j g_{ki} - g^{ki}\partial_i g_{jk}) \qquad \text{(symmetry of metric in 3}^{\text{rd}}\text{ term)} \\
&= \tfrac{1}{2}\, g^{ki}\partial_j g_{ki}.
\end{aligned}$$

Thus, equation (88) becomes:

$$\nabla_i A^i \;=\; \partial_i A^i + \frac{A^i}{2} g^{jk}\partial_i g_{jk}, \tag{89}$$

where the dummy indices have been renamed in the last term, which is a *triple* sum. We will simplify this triple sum considerably by what may seem to be a rather circuitous route.

Theorem 5.1. (Jacobi's Formula) *Let Q_{ij} be a rank 2 covariant tensor of dimension m (and thus can be represented by an $m \times m$ matrix), and let $Q = \det(Q_{ij})$ be its determinant. Let*

$\mathcal{Q}^{ji}$ be the cofactor of Q_{ij}[11]*, and thus as a matrix, $\mathcal{Q}^{ji}$ is the transpose of the cofactors of Q_{ij}*[12]*. Then,*

$$\partial_k Q = \mathcal{Q}^{ji}\partial_k Q_{ij} \qquad \textit{Jacobi's formula,} \tag{90}$$

where the double sum yields the trace of the product of the matrices $\mathcal{Q}^{ji}$ and $\partial_k Q_{ij}$.

Proof. Q can be thought of as a function of the matrix elements, Q_{ij}, and thus we have by the chain rule:

$$\partial_k Q = \frac{\partial Q}{\partial Q_{ij}}\partial_k Q_{ij}, \tag{91}$$

a double sum. Now, Laplace's formula for computing the determinant is:

$$Q = \sum_{k=1}^{m} Q_{ik}\mathcal{Q}^{ki} \qquad \text{for any } i = 1, \dots, m \text{ (no sum on } i\text{)}$$

$$\Rightarrow \quad \frac{\partial Q}{\partial Q_{ij}} = \frac{\partial}{\partial Q_{ij}}\sum_{k=1}^{m} Q_{ik}\mathcal{Q}^{ki} = \sum_{k=1}^{m}\left(\frac{\partial Q_{ik}}{\partial Q_{ij}}\mathcal{Q}^{ki} + Q_{ik}\frac{\partial \mathcal{Q}^{ki}}{\partial Q_{ij}}\right). \tag{92}$$

Now, $\partial Q_{ik}/\partial Q_{ij} = \delta_k{}^j$ since the matrix elements are independent of each other. Further, $\partial\mathcal{Q}^{ki}/\partial Q_{ij} = 0$ since the cofactor $\mathcal{Q}^{ki}$ includes all matrix elements other than those in column i and row k and must therefore be independent of Q_{ij} which is an element in the i^{th} column. Thus, equation (92) simplifies to:

$$\frac{\partial Q}{\partial Q_{ij}} = \sum_{k=1}^{m}\delta_k{}^j\mathcal{Q}^{ki} = \mathcal{Q}^{ji}.$$

Substituting this result into equation (91) gives us our desired result. □

Now, if $Q_{ij} = g_{ij}$, the metric, equation (90) becomes:

$$\partial_k g = \mathcal{G}^{ji}\partial_k g_{ij}, \tag{93}$$

where $g = \det(g_{ij})$ and $\mathcal{G}^{ji}$ is the cofactor of g_{ij}. By a variation of Cramer's rule, the inverse of g_{ij}, let us write it as g^{ji} (equation 20), is given by:

$$g^{ji} = \frac{1}{g}\mathcal{G}^{ji} \quad \Rightarrow \quad \mathcal{G}^{ji} = g\,g^{ij}$$

since g^{ij} is symmetric. Substituting this into equation (93) gives us:

$$\partial_k g = g\,g^{ij}\partial_k g_{ij} \quad \Rightarrow \quad g^{jk}\partial_i g_{jk} = \frac{1}{g}\,\partial_i g,$$

[11] In matrix algebra, the cofactor of the ij^{th} matrix element is what one gets when the i^{th} column and j^{th} row are struck out, and the resulting $m-1 \times m-1$ matrix is multiplied by -1^{i+j}.

[12] $\mathcal{Q}^{ji}$ is known as the *adjugate* of Q_{ij}.

taking the liberty once again to rename the indices. We substitute this result into equation (89) to get:

$$\nabla_i A^i \;=\; \partial_i A^i + A^i \frac{1}{2g}\,\partial_i g \;=\; \partial_i A^i + A^i \frac{1}{\sqrt{g}}\,\partial_i \sqrt{g} \;=\; \frac{1}{\sqrt{g}}\,\partial_i(\sqrt{g}A^i). \tag{94}$$

The final equality is known as the *Voss-Weyl formula*, and finally brings us to the purpose of this subsection. For orthogonal coordinates only, g_{ij} is diagonal and $\sqrt{g} = h_{(1)}h_{(2)}h_{(3)}$. Thus, with the aid of equation (29), equation (94) becomes:

$$\begin{aligned}\nabla_i A^i \;&=\; \frac{1}{h_{(1)}h_{(2)}h_{(3)}}\partial_i(h_{(1)}h_{(2)}h_{(3)}A^i)\\ &=\; \frac{1}{h_{(1)}h_{(2)}h_{(3)}}\sum_{i=1}^{3}\partial_i\left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}}A_{(i)}\right) \;=\; \nabla\cdot\vec{A},\end{aligned} \tag{95}$$

recovering the vector calculus definition of the vector divergence in orthogonal coordinates.

5.3 Divergence of a tensor

Definition 5.1. The *divergence of a contravariant tensor*, T, is the contraction of the covariant derivative with the *first* index of the tensor, and is itself a contravariant tensor of rank one less than T. Specifically, for a rank 2 tensor, we have:

$$(\nabla\cdot\mathsf{T})^j \;\equiv\; \nabla_i T^{ij}, \tag{96}$$

with similar expressions applying for tensors of higher rank.[13]

Thus, the physical component of the tensor divergence is, by equation (29),

$$(\nabla\cdot\mathsf{T})_{(j)} \;=\; h_{(j)}(\nabla\cdot\mathsf{T})^j \;=\; h_{(j)}\nabla_i T^{ij}. \tag{97}$$

Now, by equation (81), we have:

$$\nabla_i T^{ij} \;=\; \partial_i T^{ij} + \Gamma^i_{ik}T^{kj} + \Gamma^j_{ik}T^{ik} \;=\; \frac{1}{\sqrt{g}}\,\partial_i(\sqrt{g}T^{ij}) + \Gamma^j_{ik}T^{ik}, \tag{98}$$

using the Voss-Weyl formula (equation 94) on the first two terms of the middle expression. For the last term, we start by examining the $j = 1$ component:

$$\begin{aligned}\Gamma^1_{ik}T^{ik} \;=\; &\Gamma^1_{11}T^{11} + \Gamma^1_{12}T^{12} + \Gamma^1_{13}T^{13} + \underbrace{\Gamma^1_{21}}_{\Gamma^1_{12}}T^{21} + \Gamma^1_{22}T^{22} + \Gamma^1_{23}T^{23} + \underbrace{\Gamma^1_{31}}_{\Gamma^1_{13}}T^{31}\\ &+ \Gamma^1_{32}T^{32} + \Gamma^1_{33}T^{33}\end{aligned}$$

[13] The definition in equation (96) agrees with Weinberg (*Gravitation and Cosmology*, ISBN 0-471-92567-5, page 107, equation 4.7.9) and disagrees with Stone & Norman, 1992, ApJS, vol. 80, page 788, where they define $(\nabla\cdot\mathsf{T})^i = \nabla_j T^{ij}$. Thus, Weinberg contracts on the first index of T (in keeping with the ordinary rules of matrix multiplication) while SN contract on the second, and these are the same *only* if T is symmetric. Further, SN seem to have a sign error on the leading term of the second line of their equation (130), and similarly the second/third leading term of the second line of their equation (131)/(132).

given the symmetry of the first two indices on the Christoffel symbols. Now from equation (32), we have $T_{(ik)} = h_{(i)}h_{(k)}T^{ik}$. Further, if we assume orthogonal coordinates (as we do for the remainder of this section), we can use equation (68) and remark 3.1 to write:

$$\Gamma^j_{jk} = \frac{1}{h_{(j)}}\partial_k h_{(j)}, \; j,k = 1,2,3; \quad \Gamma^j_{kk} = -\frac{h_{(k)}}{h^2_{(j)}}\partial_j h_{(k)}, \; j \neq k; \quad \Gamma^j_{ik} = 0, \; i \neq j \neq k.$$

Whence,

$$\begin{aligned}
\Gamma^1_{ik}T^{ik} &= \frac{\partial_1 h_{(1)}}{h_{(1)}}\frac{T_{(11)}}{h^2_{(1)}} + \frac{\partial_2 h_{(1)}}{h_{(1)}}\frac{T_{(12)}}{h_{(1)}h_{(2)}} + \frac{\partial_3 h_{(1)}}{h_{(1)}}\frac{T_{(13)}}{h_{(1)}h_{(3)}} + \frac{\partial_2 h_{(1)}}{h_{(1)}}\frac{T_{(21)}}{h_{(2)}h_{(1)}} \\
&\quad - \frac{h_{(2)}}{h^2_{(1)}}\partial_1 h_{(2)}\frac{T_{(22)}}{h^2_{(2)}} + \frac{\partial_3 h_{(1)}}{h_{(1)}}\frac{T_{(31)}}{h_{(3)}h_{(1)}} - \frac{h_{(3)}}{h^2_{(1)}}\partial_1 h_{(3)}\frac{T_{(33)}}{h^2_{(3)}} \\
&= \frac{1}{h^2_{(1)}}\Bigg(\frac{T_{(11)}}{h_{(1)}}\partial_1 h_{(1)} - \frac{T_{(22)}}{h_{(2)}}\partial_1 h_{(2)} - \frac{T_{(33)}}{h_{(3)}}\partial_1 h_{(3)} \\
&\qquad + \frac{T_{(12)} + T_{(21)}}{h_{(2)}}\partial_2 h_{(1)} + \frac{T_{(13)} + T_{(31)}}{h_{(3)}}\partial_3 h_{(1)}\Bigg),
\end{aligned}$$

which, for any j, may be written:

$$\Gamma^j_{ik}T^{ik} = \sum_i \frac{1}{h^2_{(j)}h_{(i)}}\Big(\big(T_{(ji)} + T_{(ij)}\big)\partial_i h_{(j)} - T_{(ii)}\partial_j h_{(i)}\Big). \tag{99}$$

This is identically zero if T is antisymmetric, as expected since Γ^j_{ik} is symmetric in i and k. Note that the dummy indices i on each side of equation (99) are unrelated.

Next, in analogy to equation (95), we write:

$$\frac{1}{\sqrt{g}}\,\partial_i(\sqrt{g}T^{ij}) = \frac{1}{h_{(1)}h_{(2)}h_{(3)}}\sum_i \partial_i\left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}h_{(j)}}T_{(ij)}\right), \tag{100}$$

again using equation (32). Substituting equations (99) and (100) into equation (98), and then substituting that result into equation (97), we get our final expression for the physical component of the divergence of a rank 2 contravariant tensor:

$$\begin{aligned}
(\nabla \cdot \mathsf{T})_{(j)} &= \frac{h_{(j)}}{h_{(1)}h_{(2)}h_{(3)}}\sum_i \partial_i\left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}h_{(j)}}T_{(ij)}\right) \\
&\quad + \sum_i \frac{1}{h_{(j)}h_{(i)}}\Big(\big(T_{(ji)} + T_{(ij)}\big)\partial_i h_{(j)} - T_{(ii)}\partial_j h_{(i)}\Big).
\end{aligned} \tag{101}$$

5.4 The Laplacian of a scalar

In vector notation, the Laplacian of a scalar, f, is given by:

$$\nabla^2 f = \nabla \cdot \nabla f,$$

and thus the Laplacian is the divergence of the vector ∇f. Now, §5.2 defines the tensor divergence as an operator acting on a *contravariant* vector, V^i, whereas §5.1 defines the gradiant as a *covariant* vector, $\partial_i f$. Thus, to turn the tensor gradient into a contavariant vector, we need to multiply it by g^{ij}, whence:

$$\nabla^2 f \;=\; \nabla_i(g^{ij}\partial_j f) \;=\; \frac{1}{\sqrt{g}}\partial_i(\sqrt{g}g^{ij}\partial_j f). \tag{102}$$

Once again, for orthogonal coordinates, $g^{ij} = h_{(i)}^{-1}h_{(j)}^{-1}\delta^{ij}$ and $\sqrt{g} = h_{(1)}h_{(2)}h_{(3)}$. Thus, equation (102) becomes:

$$\nabla^2 f \;=\; \frac{1}{h_{(1)}h_{(2)}h_{(3)}}\sum_{i=1}^{3}\partial_i\left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}^2}\partial_i f\right). \tag{103}$$

5.5 Curl of a vector

Consider the following difference of covariant derivatives of a covariant vector:

$$\nabla_j A_i - \nabla_i A_j \;=\; \partial_j A_i - \Gamma_{ij}^k A_k - \partial_i A_j + \Gamma_{ji}^k A_k \;=\; \partial_j A_i - \partial_i A_j,$$

because of the symmetry of the lower indices on the Christoffel symbol. Thus, $\partial_j A_i - \partial_i A_j$ forms a tensor. By contrast, note that in

$$\nabla_j A^i - \nabla_i A^j \;=\; \partial_j A^i + \Gamma_{jk}^i A^k - \partial_i A^j - \Gamma_{ik}^j A^k,$$

the Christoffel symbols do not cancel, and thus $\partial_j A^i - \partial_i A^j$ does not form a tensor.

Now, while the construct $\partial_j A_i - \partial_i A_j$ is highly reminiscent of a curl, it cannot be what we seek since the curl is a vector (rank 1 tensor) with just one index, while $\partial_j A_i - \partial_i A_j$ is a rank 2 tensor with two indices. To form the tensor curl, we use the permutation tensor (equation 58) in a similar manner to how cross products are constructed from the Levi-Civita symbol (equation 54).

Definition 5.2. Given a rank 1 covariant tensor, A_j,

$$C^k \;\equiv\; \epsilon^{ijk}\partial_i A_j \;=\; \frac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_i A_j, \tag{104}$$

is the contravariant rank 1 *tensor curl*.

By inspection, one can easily verify that in the implied double sum, all partial derivatives of A_j appear in differences, each being the component of a tensor by the opening argument. Since ϵ^{ijk} is a component of a rank 3 contravariant tensor, then C^k must be a tensor too. Thus, tensor curls are rank 1 contravariant vectors created from rank 1 covariant vectors.

To get the physical curl, we use equation (29) to get $C^k = C_{(k)}/h_{(k)}$ (valid for all coordinate systems) and equation (31) to get $A_j = A_{(j)}h_{(j)}$ (valid for orthogonal coordinates only), and substitute these into equation (104) to get:

$$C_{(k)} \;=\; \frac{h_{(k)}}{\sqrt{g}}\sum_{ij}\varepsilon^{ijk}\partial_i(h_{(j)}A_{(j)})$$

$$\begin{aligned}\Rightarrow \quad \vec{C} \;=\; \Bigg(&\frac{1}{h_{(2)}h_{(3)}}\Big(\partial_2(h_{(3)}A_{(3)}) - \partial_3(h_{(2)}A_{(2)})\Big),\\ &\frac{1}{h_{(3)}h_{(1)}}\Big(\partial_3(h_{(1)}A_{(1)}) - \partial_1(h_{(3)}A_{(3)})\Big),\\ &\frac{1}{h_{(1)}h_{(2)}}\Big(\partial_1(h_{(2)}A_{(2)}) - \partial_2(h_{(1)}A_{(1)})\Big)\Bigg) \;=\; \nabla\times\vec{A},\end{aligned} \tag{105}$$

the vector calculus definition of the curl for orthogonal coordinates.

5.6 The Laplacian of a vector

The Laplacian of a vector, $\nabla^2\vec{A}$, can be most easily written down in invariant form by using the identity:

$$\nabla^2\vec{A} \;=\; \nabla(\nabla\cdot\vec{A}) - \nabla\times(\nabla\times\vec{A}), \tag{106}$$

and using the invariant expansions for each of the gradient of a scalar (equation 87), divergence of a vector (equation 95), and the curl (equation 105).

5.7 Gradient of a vector

The covariant derivative of a covariant vector, known in vector calculus as the *vector gradient*, is given by equation (80), reproduced below for convenience:

$$\nabla_i A_j \;=\; \partial_i A_j - \Gamma^k_{ij}A_k \;\equiv\; G_{ij},$$

where G_{ij} is a rank 2 covariant tensor. To express this in terms of physical components, we restrict ourselves once again to orthogonal coordinates, where equation (31) $\Rightarrow A_j = h_{(j)}A_{(j)}$ and equation (33) $\Rightarrow G_{ij} = h_{(i)}h_{(j)}G_{(ij)}$. Substituting these into equation (80), we have:

$$G_{(ij)} \;=\; \frac{1}{h_{(i)}h_{(j)}}\left(\partial_i(h_{(j)}A_{(j)}) - \sum_k \Gamma^k_{ij}h_{(k)}A_{(k)}\right) \;\equiv\; (\nabla\vec{A})_{(ij)}. \tag{107}$$

Now, for orthogonal coordinates, the Christoffel symbols simplify considerably. Substituting the results of equation (68) into equation (107), we get for the (11) component:

$$\begin{aligned}(\nabla\vec{A})_{(11)} \;&=\; \frac{1}{h^2_{(1)}}\big(\partial_1(h_{(1)}A_{(1)}) - \Gamma^1_{11}h_{(1)}A_{(1)} - \Gamma^2_{11}h_{(2)}A_{(2)} - \Gamma^3_{11}h_{(3)}A_{(3)}\big)\\ &=\; \frac{1}{h^2_{(1)}}\Bigg(h_{(1)}\partial_1A_{(1)} + A_{(1)}\partial_1h_{(1)}\\ &\qquad -\; A_{(1)}h_{(1)}\frac{1}{2g_{11}}\partial_1g_{11} + A_{(2)}h_{(2)}\frac{1}{2g_{22}}\partial_2g_{11} + A_{(3)}h_{(3)}\frac{1}{2g_{33}}\partial_3g_{11}\Bigg)\\ &=\; \frac{1}{h^2_{(1)}}\left(h_{(1)}\partial_1A_{(1)} + A_{(1)}\partial_1h_{(1)} - A_{(1)}\partial_1h_{(1)} + A_{(2)}\frac{h_{(1)}}{h_{(2)}}\partial_2h_{(1)} + A_{(3)}\frac{h_{(1)}}{h_{(3)}}\partial_3h_{(1)}\right)\\ &=\; \frac{h_{(1)}\partial_1A_{(1)} - A_{(1)}\partial_1h_{(1)}}{h^2_{(1)}} + \frac{1}{h_{(1)}}\left(\frac{A_{(1)}}{h_{(1)}}\partial_1h_{(1)} + \frac{A_{(2)}}{h_{(2)}}\partial_2h_{(1)} + \frac{A_{(3)}}{h_{(3)}}\partial_3h_{(1)}\right)\end{aligned}$$

$$= \partial_1\left(\frac{A_{(1)}}{h_{(1)}}\right) + \frac{1}{h_{(1)}}\vec{A}\cdot\nabla h_{(1)}.$$

For the (12) component, we get:

$$\begin{aligned}
(\nabla\vec{A})_{(12)} &= \frac{1}{h_{(1)}h_{(2)}}\left(\partial_1(h_{(2)}A_{(2)}) - \Gamma^1_{12}h_{(1)}A_{(1)} - \Gamma^2_{12}h_{(2)}A_{(2)} - \Gamma^3_{12}h_{(3)}A_{(3)}\right) \\
&= \frac{1}{h_{(1)}h_{(2)}}\left(h_{(2)}\partial_1 A_{(2)} + A_{(2)}\partial_1 h_{(2)} - A_{(1)}h_{(1)}\frac{1}{2g_{11}}\partial_2 g_{11} - A_{(2)}h_{(2)}\frac{1}{2g_{22}}\partial_1 g_{22}\right) \\
&= \frac{1}{h_{(1)}h_{(2)}}\left(h_{(2)}\partial_1 A_{(2)} + \cancel{A_{(2)}\partial_1 h_{(2)}} - A_{(1)}\partial_2 h_{(1)} - \cancel{A_{(2)}\partial_1 h_{(2)}}\right) \\
&= \frac{1}{h_{(1)}}\left(\partial_1 A_{(2)} - \frac{A_{(1)}}{h_{(2)}}\partial_2 h_{(1)}\right),
\end{aligned}$$

Evidently, all components of $\nabla\vec{A}$ may be written as:

$$(\nabla\vec{A})_{(ij)} = \begin{cases} \partial_i\left(\dfrac{A_{(i)}}{h_{(i)}}\right) + \dfrac{1}{h_{(i)}}\vec{A}\cdot\nabla h_{(i)}, & i = j; \\ \dfrac{1}{h_{(i)}}\left(\partial_i A_{(j)} - \dfrac{A_{(i)}}{h_{(j)}}\partial_j h_{(i)}\right), & i \neq j. \end{cases} \tag{108}$$

5.8 Summary

For ease of reference, the main results of this section are gathered with their equation numbers. If f, $\vec{A}$, and T are arbitrary scalar, vector, and contravariant rank 2 tensor functions of the *orthogonal* coordinates, we have found:

$$\nabla f = \left(\frac{1}{h_{(1)}}\partial_1 f,\ \frac{1}{h_{(2)}}\partial_2 f,\ \frac{1}{h_{(3)}}\partial_3 f\right); \tag{87}$$

$$\nabla^2 f = \frac{1}{h_{(1)}h_{(2)}h_{(3)}}\sum_i \partial_i\left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h^2_{(i)}}\partial_i f\right); \tag{103}$$

$$\nabla\cdot\vec{A} = \frac{1}{h_{(1)}h_{(2)}h_{(3)}}\sum_i \partial_i\left(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}}A_{(i)}\right); \tag{95}$$

$$\begin{aligned}
\nabla\times\vec{A} = \Bigg(&\frac{1}{h_{(2)}h_{(3)}}\Big(\partial_2(h_{(3)}A_{(3)}) - \partial_3(h_{(2)}A_{(2)})\Big), \frac{1}{h_{(3)}h_{(1)}}\Big(\partial_3(h_{(1)}A_{(1)}) - \partial_1(h_{(3)}A_{(3)})\Big), \\
&\frac{1}{h_{(1)}h_{(2)}}\Big(\partial_1(h_{(2)}A_{(2)}) - \partial_2(h_{(1)}A_{(1)})\Big)\Bigg);
\end{aligned} \tag{105}$$

$$(\nabla\vec{A})_{(ij)} = \begin{cases} \partial_i\left(\dfrac{A_{(i)}}{h_{(i)}}\right) + \dfrac{1}{h_{(i)}}\vec{A}\cdot\nabla h_{(i)}, & i = j, \\ \dfrac{1}{h_{(i)}}\left(\partial_i A_{(j)} - \dfrac{A_{(i)}}{h_{(j)}}\partial_j h_{(i)}\right), & i \neq j; \end{cases} \tag{108}$$

$$
\begin{aligned}
(\nabla\cdot\mathsf{T})_{(j)} \;=\; & \frac{h_{(j)}}{h_{(1)}h_{(2)}h_{(3)}}\sum_i \partial_i\Big(\frac{h_{(1)}h_{(2)}h_{(3)}}{h_{(i)}h_{(j)}}T_{(ij)}\Big) \\
& + \sum_i \frac{1}{h_{(j)}h_{(i)}}\Big(\big(T_{(ji)}+T_{(ij)}\big)\partial_i h_{(j)} - T_{(ii)}\partial_j h_{(i)}\Big).
\end{aligned}
\tag{101}
$$

For non-orthogonal coordinates, these expressions become rather more cumbersome.

5.9 A tensor-vector identity

A useful relation in vector calculus, particularly for extending the ideal fluid equations to include viscous stresses, is the following:

Theorem 5.2. *If T is a rank 2 tensor and $\vec{A}$ is a vector (rank 1 tensor), then, in terms of their physical components:*

$$\nabla\cdot(\mathsf{T}\cdot\vec{A}) \;=\; \mathsf{T}:\nabla\vec{A} + (\nabla\cdot\mathsf{T})\cdot\vec{A}. \tag{109}$$

Implicit in this identity are the definitions of the "dot" product between two vectors, the "dot" product between a vector and a rank 2 tensor, and the "colon" product between two rank 2 tensors (§3.2). With these in hand, the proof of the theorem is a simple matter of bringing together the relevant bits from this primer.

Proof. Start with Theorem 4.4, the product rule for covariant differentiation. Thus, if $B^i = T^{ij}A_j$, then,

$$\nabla_i B^i \;=\; \nabla_i(T^{ij}A_j) \;=\; T^{ij}\nabla_i A_j + A_j\nabla_i T^{ij}. \tag{110}$$

Now, from equation (95), $\nabla_i B^i = \nabla\cdot\vec{B}$, where $\vec{B}$ is the ordered triple of *physical* components. From equation (42), we have $B^i = T^{ij}A_j = (\mathsf{T}\cdot\vec{A})^i$ (the right dot product), and thus:

$$\nabla_i B^i \;=\; \nabla\cdot(\mathsf{T}\cdot\vec{A}). \tag{111}$$

Next, T^{ij} is a rank 2 contravariant tensor, $\nabla_i A_j$ is a rank 2 covariant tensor and, according to equation (41),

$$T^{ij}(\nabla_i A_j) \;=\; \mathsf{T}:\nabla\vec{A}. \tag{112}$$

Finally, according to equation (96), $\nabla_i T^{ij} = (\nabla\cdot T)^j$ and, by equation (40), $A_j(\nabla\cdot T)^j = \vec{A}\cdot(\nabla\cdot\mathsf{T})$. Thus,

$$A_j\nabla_i T^{ij} \;=\; \vec{A}\cdot(\nabla\cdot\mathsf{T}). \tag{113}$$

Note the right hand sides of equations (111), (112), and (113) are all in terms of physical components. Substituting these equations into equation (110) proves the theorem. □

For orthogonal coordinates, we can confirm this identity directly using the results from §5.8. Start with the LHS of equation (109) by substituting equation (45) into equation (95):

$$\nabla\cdot(\mathsf{T}\cdot\vec{A}) \;=\; \frac{1}{h_1h_2h_3}\sum_i\partial_i\Big(\frac{h_1h_2h_3}{h_i}\sum_j T_{ij}A_j\Big) \;=\; \frac{1}{h_1h_2h_3}\sum_{ij}\partial_i\Big(\frac{h_1h_2h_3}{h_i}T_{ij}A_j\Big). \tag{114}$$

With the understanding that all components are physical components, we drop the parentheses on the subscripts for convenience.

Now, using equation (41) we write:

$$
\begin{aligned}
T : \nabla\vec{A} = \sum_{ij} T_{ij}(\nabla\vec{A})_{ij} &= \sum_{i} T_{ii}(\nabla\vec{A})_{ii} + \sum_{i\neq j} T_{ij}(\nabla\vec{A})_{ij} \\
&= \sum_{i}\left(T_{ii}\partial_i\left(\frac{A_i}{h_i}\right) + \frac{T_{ii}}{h_i}\sum_{j}\frac{A_j}{h_j}\partial_j h_i\right) + \sum_{i\neq j}\left(\frac{T_{ij}}{h_i}\partial_i A_j - \frac{T_{ij}A_i}{h_i h_j}\partial_j h_i\right),
\end{aligned}
$$

using equation (108). Then, from equation (101) we have:

$$
(\nabla\cdot\mathsf{T})\cdot\vec{A} = \sum_{j}\frac{A_j h_j}{h_1 h_2 h_3}\sum_{i}\partial_i\left(\frac{h_1 h_2 h_3}{h_i h_j}T_{ij}\right) + \sum_{ij}\frac{A_j}{h_j h_i}\Big((T_{ji}+T_{ij})\partial_i h_j - T_{ii}\partial_j h_i\Big).
$$

Thus, the RHS of equation (109) becomes:

$$
\begin{aligned}
\mathsf{T}:\nabla\vec{A} + (\nabla\cdot\mathsf{T})\cdot\vec{A} &= \sum_{i}\frac{T_{ii}}{h_i}\partial_i A_i - \sum_{i}\frac{T_{ii}A_i}{h_i^2}\partial_i h_i + \cancel{\sum_{ij}\frac{T_{ii}A_j}{h_i h_j}\partial_j h_i} + \sum_{i\neq j}\frac{T_{ij}}{h_i}\partial_i A_j \\
&\quad - \sum_{i\neq j}\frac{T_{ij}A_i}{h_i h_j}\partial_j h_i + \sum_{ij}\frac{A_j h_j}{h_1 h_2 h_3}\partial_i\left(\frac{h_1 h_2 h_3}{h_i h_j}T_{ij}\right) + \underbrace{\sum_{ij}\frac{A_j T_{ji}}{h_j h_i}\partial_i h_j}_{\text{swap } i \text{ and } j} \\
&\quad + \sum_{ij}\frac{A_j T_{ij}}{h_j h_i}\partial_i h_j - \cancel{\sum_{ij}\frac{A_j T_{ii}}{h_j h_i}\partial_j h_i} \\
&= \underbrace{\sum_{i}\frac{T_{ii}}{h_i}\partial_i A_i + \sum_{i\neq j}\frac{T_{ij}}{h_i}\partial_i A_j}_{\sum_{ij}\frac{T_{ij}}{h_i}\partial_i A_j} \underbrace{- \sum_{i}\frac{T_{ii}A_i}{h_i^2}\partial_i h_i - \sum_{i\neq j}\frac{T_{ij}A_i}{h_i h_j}\partial_j h_i}_{-\cancel{\sum_{ij}\frac{T_{ij}A_i}{h_i h_j}\partial_j h_i}} \\
&\quad + \sum_{ij}\frac{A_j h_j}{h_1 h_2 h_3}\partial_i\left(\frac{h_1 h_2 h_3}{h_i h_j}T_{ij}\right) + \cancel{\sum_{ij}\frac{A_i T_{ij}}{h_i h_j}\partial_j h_i} + \sum_{ij}\frac{A_j T_{ij}}{h_j h_i}\partial_i h_j \\
&= \sum_{ij}\left[\frac{A_j h_j}{h_1 h_2 h_3}\partial_i\left(\frac{h_1 h_2 h_3}{h_i h_j}T_{ij}\right) + \underbrace{\frac{T_{ij}}{h_i}\partial_i A_j + \frac{A_j T_{ij}}{h_j h_i}\partial_i h_j}_{\frac{T_{ij}}{h_i h_j}\partial_i(A_j h_j)}\right] \\
&= \frac{1}{h_1 h_2 h_3}\sum_{ij}\left[A_j h_j\partial_i\left(\frac{h_1 h_2 h_3}{h_i h_j}T_{ij}\right) + \frac{h_1 h_2 h_3}{h_i h_j}T_{ij}\,\partial_i(A_j h_j)\right] \\
&= \frac{1}{h_1 h_2 h_3}\sum_{ij}\partial_i\left(\frac{h_1 h_2 h_3}{h_i \cancel{h_j}}T_{ij}A_j\cancel{h_j}\right) \;=\; \text{LHS (equation 114).}
\end{aligned}
$$

6 Cartesian, cylindrical, spherical polar coordinates

Table 1 in the beginning of §3 gives the scale factors, $h_{(i)}$, for each of Cartesian, (x, y, z), cylindrical, (z, ϱ, φ), and spherical polar, (r, ϑ, φ), coordinates. Restricted to these coordinate systems, h_3 is a separable function of its arguments, and we can write $h_3(x_1, x_2) \equiv h_{31}(x_1)h_{32}(x_2)$, dropping the parentheses from all subscripts now that everything is being expressed in terms of the *physical* components. The scaling factors then become:

$$
\begin{aligned}
h_1 \;&=\; 1, && \text{all;} \\
h_2(x_1) \;=\; h_{31}(x_1) \;&=\; \begin{cases} 1, & \text{Cartesian and cylindrical,} \\ x_1 = r, & \text{spherical polar;} \end{cases} \\
h_{32}(x_2) \;&=\; \begin{cases} 1, & \text{Cartesian,} \\ x_2 = \varrho, & \text{cylindrical,} \\ \sin x_2 = \sin\vartheta, & \text{spherical polar.} \end{cases}
\end{aligned}
$$

With this, the six tensor constructs in §5.8 can then be written in their full gory detail as:

$$\nabla f \;=\; \left(\partial_1 f,\, \frac{1}{h_2}\partial_2 f,\, \frac{1}{h_3}\partial_3 f\right); \tag{115}$$

$$\nabla^2 f \;=\; \frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31}\partial_1 f) + \frac{1}{h_2^2 h_{32}}\partial_2(h_{32}\partial_2 f) + \frac{1}{h_3^2}\partial_3^2 f; \tag{116}$$

$$\nabla\cdot\vec{A} \;=\; \frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} A_1) + \frac{1}{h_2 h_{32}}\partial_2(h_{32} A_2) + \frac{1}{h_3}\partial_3 A_3; \tag{117}$$

$$\nabla\times\vec{A} \;=\; \left(\frac{1}{h_2 h_{32}}\partial_2(h_{32}A_3) - \frac{1}{h_3}\partial_3 A_2,\, \frac{1}{h_3}\partial_3 A_1 - \frac{1}{h_{31}}\partial_1(h_{31}A_3),\, \frac{1}{h_2}\big(\partial_1(h_2 A_2) - \partial_2 A_1\big)\right); \tag{118}$$

$$\nabla\vec{A} \;=\; \begin{bmatrix} \partial_1 A_1 & \partial_1 A_2 & \partial_1 A_3 \\ \frac{1}{h_2}\big(\partial_2 A_1 - A_2\partial_1 h_2\big) & \frac{1}{h_2}\big(\partial_2 A_2 + A_1\partial_1 h_2\big) & \frac{1}{h_2}\partial_2 A_3 \\ \frac{1}{h_3}\partial_3 A_1 - \frac{A_3}{h_{31}}\partial_1 h_{31} & \frac{1}{h_3}\partial_3 A_2 - \frac{A_3}{h_2 h_{32}}\partial_2 h_{32} & \frac{1}{h_3}\partial_3 A_3 + \frac{A_1}{h_{31}}\partial_1 h_{31} + \frac{A_2}{h_2 h_{32}}\partial_2 h_{32} \end{bmatrix}; \tag{119}$$

$$
\begin{aligned}
\nabla\cdot\mathsf{T} \;=\; \Bigg(&\frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} T_{11}) + \frac{1}{h_2 h_{32}}\partial_2(h_{32} T_{21}) + \frac{1}{h_3}\partial_3 T_{31} - \frac{T_{22}}{h_2}\partial_1 h_2 - \frac{T_{33}}{h_{31}}\partial_1 h_{31}, \\
&\frac{1}{h_{31}}\partial_1(h_{31} T_{12}) + \frac{1}{h_2 h_{32}}\partial_2(h_{32} T_{22}) + \frac{1}{h_3}\partial_3 T_{32} + \frac{T_{21}+T_{12}}{h_2}\partial_1 h_2 - \frac{T_{33}}{h_2 h_{32}}\partial_2 h_{32}, \\
&\frac{1}{h_2}\partial_1(h_2 T_{13}) + \frac{1}{h_2}\partial_2 T_{23} + \frac{1}{h_3}\partial_3 T_{33} + \frac{T_{31}+T_{13}}{h_{31}}\partial_1 h_{31} + \frac{T_{32}+T_{23}}{h_2 h_{32}}\partial_2 h_{32}\Bigg).
\end{aligned} \tag{120}
$$

6.1 Cartesian coordinates

In Cartesian coordinates, equations (115) – (120) become:

$$\nabla f = \big(\partial_x f, \partial_y f, \partial_z f\big);$$

$$\nabla^2 f = \partial_x^2 f + \partial_y^2 f + \partial_z^2 f;$$

$$\nabla \cdot \vec{A} = \partial_x A_x + \partial_y A_y + \partial_z A_z;$$

$$\nabla \times \vec{A} = \big(\partial_y A_z - \partial_z A_y,\, \partial_z A_x - \partial_x A_z,\, \partial_x A_y - \partial_y A_x\big);$$

$$\nabla \vec{A} = \begin{bmatrix} \partial_x A_x & \partial_x A_y & \partial_x A_z \\ \partial_y A_x & \partial_y A_y & \partial_y A_z \\ \partial_z A_x & \partial_z A_y & \partial_z A_z \end{bmatrix};$$

$$\nabla \cdot \mathsf{T} = \big(\partial_x T_{xx} + \partial_y T_{yx} + \partial_z T_{zx},\ \partial_x T_{xy} + \partial_y T_{yy} + \partial_z T_{zy},\ \partial_x T_{xz} + \partial_y T_{yz} + \partial_z T_{zz}\big).$$

6.2 Cylindrical coordinates

In cylindrical coordinates, equations (115) – (120) become:

$$\nabla f = \left(\partial_z f,\, \partial_\varrho f,\, \frac{1}{\varrho}\partial_\varphi f\right);$$

$$\nabla^2 f = \partial_z^2 f + \frac{1}{\varrho}\partial_\varrho(\varrho \partial_\varrho f) + \frac{1}{\varrho^2}\partial_\varphi^2 f;$$

$$\nabla \cdot \vec{A} = \partial_z A_z + \frac{1}{\varrho}\partial_\varrho(\varrho A_\varrho) + \frac{1}{\varrho}\partial_\varphi A_\varphi;$$

$$\nabla \times \vec{A} = \left(\frac{1}{\varrho}\big(\partial_\varrho(\varrho A_\varphi) - \partial_\varphi A_\varrho\big),\, \frac{1}{\varrho}\partial_\varphi A_z - \partial_z A_\varphi,\, \partial_z A_\varrho - \partial_\varrho A_z\right);$$

$$\nabla \vec{A} = \begin{bmatrix} \partial_z A_z & \partial_z A_\varrho & \partial_z A_\varphi \\ \partial_\varrho A_z & \partial_\varrho A_\varrho & \partial_\varrho A_\varphi \\ \frac{1}{\varrho}\partial_\varphi A_z & \frac{1}{\varrho}\big(\partial_\varphi A_\varrho - A_\varphi\big) & \frac{1}{\varrho}\big(\partial_\varphi A_\varphi + A_\varrho\big) \end{bmatrix};$$

$$\nabla \cdot \mathsf{T} = \Bigg(\partial_z T_{zz} + \frac{1}{\varrho}\partial_\varrho(\varrho T_{\varrho z}) + \frac{1}{\varrho}\partial_\varphi T_{\varphi z},\ \partial_z T_{z\varrho} + \frac{1}{\varrho}\partial_\varrho(\varrho T_{\varrho\varrho}) + \frac{1}{\varrho}\partial_\varphi T_{\varphi\varrho} - \frac{T_{\varphi\varphi}}{\varrho},$$
$$\partial_z T_{z\varphi} + \partial_\varrho T_{\varrho\varphi} + \frac{1}{\varrho}\partial_\varphi T_{\varphi\varphi} + \frac{T_{\varphi\varrho} + T_{\varrho\varphi}}{\varrho}\Bigg).$$

6.3 Spherical polar coordinates

In spherical polar coordinates, equations (115) – (120) become:

$$\nabla f \;=\; \left(\partial_r f,\, \frac{1}{r}\partial_\vartheta f,\, \frac{1}{r\sin\vartheta}\partial_\varphi f\right);$$

$$\nabla^2 f \;=\; \frac{1}{r^2}\partial_r(r^2\partial_r f) + \frac{1}{r^2\sin\vartheta}\partial_\vartheta(\sin\vartheta\,\partial_\vartheta f) + \frac{1}{r^2\sin^2\vartheta}\partial_\varphi^2 f;$$

$$\nabla\cdot\vec{A} \;=\; \frac{1}{r^2}\partial_r(r^2 A_r) + \frac{1}{r\sin\vartheta}\partial_\vartheta(\sin\vartheta A_\vartheta) + \frac{1}{r\sin\vartheta}\partial_\varphi A_\varphi;$$

$$\nabla\times\vec{A} \;=\; \left(\frac{1}{r\sin\vartheta}\big(\partial_\vartheta(\sin\vartheta A_\varphi) - \partial_\varphi A_\vartheta\big),\, \frac{1}{r\sin\vartheta}\partial_\varphi A_r - \frac{1}{r}\partial_r(rA_\varphi),\, \frac{1}{r}\big(\partial_r(rA_\vartheta) - \partial_\vartheta A_r\big)\right);$$

$$\nabla\vec{A} \;=\; \begin{bmatrix} \partial_r A_r & \partial_r A_\vartheta & \partial_r A_\varphi \\ \frac{1}{r}\big(\partial_\vartheta A_r - A_\vartheta\big) & \frac{1}{r}\big(\partial_\vartheta A_\vartheta + A_r\big) & \frac{1}{r}\partial_\vartheta A_\varphi \\ \frac{1}{r\sin\vartheta}\partial_\varphi A_r - \frac{A_\varphi}{r} & \frac{1}{r\sin\vartheta}\partial_\varphi A_\vartheta - \frac{A_\varphi}{r\tan\vartheta} & \frac{1}{r\sin\vartheta}\partial_\varphi A_\varphi + \frac{A_r}{r} + \frac{A_\vartheta}{r\tan\vartheta} \end{bmatrix};$$

$$\begin{aligned} \nabla\cdot\mathsf{T} \;=\; & \left(\frac{1}{r^2}\partial_r(r^2 T_{rr}) + \frac{1}{r\sin\vartheta}\partial_\vartheta(\sin\vartheta\, T_{\vartheta r}) + \frac{1}{r\sin\vartheta}\partial_\varphi T_{\varphi r} - \frac{T_{\vartheta\vartheta} + T_{\varphi\varphi}}{r},\right. \\ & \frac{1}{r}\partial_r(rT_{r\vartheta}) + \frac{1}{r\sin\vartheta}\partial_\vartheta(\sin\vartheta\, T_{\vartheta\vartheta}) + \frac{1}{r\sin\vartheta}\partial_\varphi T_{\varphi\vartheta} + \frac{T_{\vartheta r} + T_{r\vartheta}}{r} - \frac{T_{\varphi\varphi}}{r\tan\vartheta}, \\ & \left.\frac{1}{r}\partial_r(rT_{r\varphi}) + \frac{1}{r}\partial_\vartheta T_{\vartheta\varphi} + \frac{1}{r\sin\vartheta}\partial_\varphi T_{\varphi\varphi} + \frac{T_{\varphi r} + T_{r\varphi}}{r} + \frac{T_{\varphi\vartheta} + T_{\vartheta\varphi}}{r\tan\vartheta}\right) \\ =\; & \left(\partial_r T_{rr} + \frac{1}{r}\partial_\vartheta T_{\vartheta r} + \frac{1}{r\sin\vartheta}\partial_\varphi T_{\varphi r} + \frac{2T_{rr} - T_{\vartheta\vartheta} - T_{\varphi\varphi}}{r} + \frac{T_{\vartheta r}}{r\tan\vartheta},\right. \\ & \partial_r T_{r\vartheta} + \frac{1}{r}\partial_\vartheta T_{\vartheta\vartheta} + \frac{1}{r\sin\vartheta}\partial_\varphi T_{\varphi\vartheta} + \frac{2T_{r\vartheta} + T_{\vartheta r}}{r} + \frac{T_{\vartheta\vartheta} - T_{\varphi\varphi}}{r\tan\vartheta}, \\ & \left.\partial_r T_{r\varphi} + \frac{1}{r}\partial_\vartheta T_{\vartheta\varphi} + \frac{1}{r\sin\vartheta}\partial_\varphi T_{\varphi\varphi} + \frac{2T_{r\varphi} + T_{\varphi r}}{r} + \frac{T_{\vartheta\varphi} + T_{\varphi\vartheta}}{r\tan\vartheta}\right). \end{aligned}$$

7 An application to viscosity

The equations of viscid hydrodynamics may be written as:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\,\vec{v}) = 0; \qquad \text{continuity}$$

$$\left.\begin{aligned} \frac{\partial \vec{s}}{\partial t} + \nabla \cdot (\vec{s}\,\vec{v}) &= -\nabla p + \nabla \cdot \mathsf{T}; && \text{momentum equation} \\ \frac{\partial \vec{v}}{\partial t} + (\vec{v}\cdot\nabla)\vec{v} &= \frac{1}{\rho}\big(-\nabla p + \nabla\cdot\mathsf{T}\big); && \text{Euler equation} \end{aligned}\right\}$$

$$\left.\begin{aligned} \frac{\partial e_{\mathrm{T}}}{\partial t} + \nabla\cdot\big((e_{\mathrm{T}} + p)\,\vec{v}\big) &= \nabla\cdot(\mathsf{T}\cdot\vec{v}); && \text{total energy equation} \\ \frac{\partial e}{\partial t} + \nabla\cdot(e\,\vec{v}) &= -p\nabla\cdot\vec{v} + \mathsf{T} : \nabla\vec{v}, && \text{internal energy equation} \end{aligned}\right\} \tag{121}$$

where one uses *one* of the momentum and Euler equations, and *one* of the total and internal energy equations. Here, ρ is the density, $\vec{v}$ is the velocity, $\vec{s} = \rho\vec{v}$ is the momentum density (and $\vec{s}\,\vec{v}$ is the dyadic product of $\vec{s}$ and $\vec{v}$; definition 2.3, equation 13), e is the internal energy density, $p = (\gamma - 1)e$ (ideal equation of state) is the thermal pressure, and $e_{\mathrm{T}} = \frac{1}{2}\rho v^2 + e$ is the total energy density. See equation (41) for a reminder of the "double dot" or "colon" product convention (*e.g.*, last term in the internal energy equation).

As for the viscid variables, T is the *viscous stress tensor* (units $\mathrm{N\,m^{-2}}$), given by:

$$\mathsf{T} = (\mu + \mu_{\mathrm{l}} + \mu_{\mathrm{q}})\,\mathsf{S}; \quad \mathsf{S} = \nabla\vec{v} + (\nabla\vec{v})^{\mathrm{T}} - \tfrac{2}{3}\nabla\cdot\vec{v}\,\mathsf{I}, \tag{122}$$

where S is the *shear tensor* (units $\mathrm{s^{-1}}$), I is the identity tensor, the superscript $^{\mathrm{T}}$ indicates the transpose, and μ is the *coefficient of shear viscosity* (units $\mathrm{N\,m^{-2}\,s}$), a physical property of the fluid. As defined, S and therefore T are symmetric.

For numerical applications in which the fluid variables are resolved on a discrete *grid* of small but not infinitesimal *zones*, most (M)HD codes are designed with an *artificial viscosity*, which helps stabilise the flow in stagnant regions and at discontinuities (shocks). These are respectively mediated by the *linear* and *quadratic* coefficients μ_{l} and μ_{q} given by:

$$\mu_{\mathrm{l}} = q_1 l \rho c_{\mathrm{s}}; \qquad \mu_{\mathrm{q}} = -q_2 l^2 \rho \min\big(0, \nabla\cdot\vec{v}\big), \tag{123}$$

where q_1 and q_2 (corresponding to $\frac{1}{2}$`qlin` and $\frac{1}{2}$`qcon` in *ZEUS*[14]) are unitless parameters (typically 0.1 and 1 respectively), l is the zone size (maximum of the three dimensions), and $c_{\mathrm{s}} = \sqrt{\gamma p/\rho}$ is the adiabatic sound speed. Note that the "min" function means μ_{q} is positive and that the quadratic viscosity is applied only in regions of compression (*e.g.*, shocks).

This form of the artificial viscosity (identical to the physical viscosity other than the coefficients) is called the *tensor artificial viscosity* and is due to W.-M. Tscharnuter & K.-H. Winkler (1979, Comput. Phys. Comm., 18, 171; TW). It differs from the more often-followed

[14] *ZEUS* and, in particular, *ZEUS-3D* is an astrophysical computational fluid dynamics code that I and others developed in the late 80s and early 90s, and which I have continued to develop since. The latest version and full documentation is freely available for download and use at http://www.ica.smu.ca/zeus3d.

approach by J. von Neumann & R. D. Richtmyer (1950, J. Appl. Phys., 21, 232; vNR) who set $\mathsf{S} = \nabla\vec{v}$, ignore all off-diagonal terms, and replace $\nabla\cdot\vec{v}$ with $\partial_i v_i$ in the i-direction for μ_{q} (equation 123).

Beginning with equation (119), the diagonal elements for $\nabla\vec{v}$ in Cartesian, cylindrical, or spherical polar coordinates, are:

$$\left.\begin{aligned}
(\nabla\vec{v})_{11} &= \partial_1 v_1;\\
(\nabla\vec{v})_{22} &= \frac{1}{h_2}\partial_2 v_2 + \frac{v_1}{h_2}\partial_1 h_2;\\
(\nabla\vec{v})_{33} &= \frac{1}{h_3}\partial_3 v_3 + \frac{v_1}{h_{31}}\partial_1 h_{31} + \frac{v_2}{h_2 h_{32}}\partial_2 h_{32}.
\end{aligned}\right\} \tag{124}$$

Note that:

$$\sum_{i=1}^{3}(\nabla\vec{v})_{ii} = \underbrace{\partial_1 v_1 + \frac{v_1}{h_2}\partial_1 h_2 + \frac{v_1}{h_{31}}\partial_1 h_{31}}_{\displaystyle\frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} v_1)} + \underbrace{\frac{1}{h_2}\partial_2 v_2 + \frac{v_2}{h_2 h_{32}}\partial_2 h_{32}}_{\displaystyle\frac{1}{h_2 h_{32}}\partial_2(h_{32} v_2)} + \frac{1}{h_3}\partial_3 v_3 = \nabla\cdot\vec{v}, \tag{125}$$

where this is a general result, true for any orthogonal coordinate system. Then, from equation (122), the diagonal elements of S are given by:

$$S_{ii} = 2\left((\nabla\vec{v})_{ii} - \frac{1}{3}\nabla\cdot\vec{v}\right), \quad i = 1, 2, 3. \tag{126}$$

Thus,

$$\mathrm{Tr}(\mathsf{S}) = \sum_{i=1}^{3} S_{ii} = 2\sum_{i=1}^{3}\left((\nabla\vec{v})_{ii} - \frac{1}{3}\nabla\cdot\vec{v}\right) = 2\sum_{i=1}^{3}(\nabla\vec{v})_{ii} - 2\sum_{i=1}^{3}\frac{1}{3}\nabla\cdot\vec{v} = 0, \tag{127}$$

because of equation (125), and both S and T are traceless.

Aside: Numerical considerations

In a numerical scheme, $\mathrm{Tr}(\mathsf{S})$ is identically zero so long as equation (125) is valid to *machine round-off error*, which does not necessarily follow from the fact that equation (125) is an algebraic identity. For those not used to the vagaries of numerical arithmetic, this inconvenient fact can come as a rude surprise.

For equation (125) to be valid numerically, expressions like:

$$\frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} v_1) = \partial_1 v_1 + \frac{v_1}{h_2}\partial_1 h_2 + \frac{v_1}{h_{31}}\partial_1 h_{31},$$

as indicated by the first underbrace must be accurate to machine round-off error. In polar coordinates where $x_1 = r$ and $h_2(r) = h_{31}(r) = r$, this means that:

$$\frac{1}{r^2}\partial_r(r^2 v_r) = \partial_r v_r + 2\frac{v_r}{r}. \tag{128}$$

Now, on the numerical grid in Fig. 2, derivatives like $\partial_r v_r$ are evaluated with *finite differences*:

$$\partial_r v_r = \frac{v_r(i+1,j) - v_r(i,j)}{\delta r},$$

which is zone-centred if v_r is face-centred. Quantities such as v_r in v_r/r that aren't differentiated are zone-centred by taking a *two-point average*:

$$\langle v_r \rangle = \frac{v_r(i+1,j) + v_r(i,j)}{2}.$$

Thus, with the exemplary values in Fig. 2, the LHS and RHS of equation (128) are evaluated as follows:

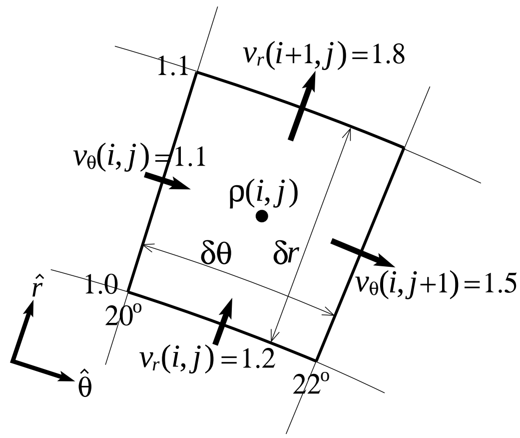

Figure 2: A single 2-D zone on a polar computational grid.

$$\begin{aligned}\text{LHS} &= \frac{1}{\langle r \rangle^2}\partial_r(r^2 v_r)\\ &= \frac{1}{(1.05)^2}\frac{(1.1)^2(1.8) - (1.0)^2(1.2)}{0.1} = 8.8707\end{aligned}$$

$$\text{RHS} = \partial_r v_r + 2\frac{\langle v_r \rangle}{\langle r \rangle} = \frac{1.8 - 1.2}{0.1} + 2\frac{1.5}{1.05} = 8.8571 \neq \text{LHS!}$$

Close, but no cigar! Similarly, equation (125) requires:

$$\begin{aligned}\frac{1}{h_2 h_{32}}\partial_2(h_{32}v_2) &= \frac{1}{h_2}\partial_2 v_2 + \frac{v_2}{h_2 h_{32}}\partial_2 h_{32}\\ \Rightarrow \quad \frac{1}{\sin\langle\theta\rangle}\partial_\theta(v_\theta \sin\theta) &= \partial_\theta v_\theta + \frac{\langle v_\theta \rangle}{\sin\langle\theta\rangle}\cos\langle\theta\rangle, \end{aligned} \tag{129}$$

since $x_2 = \theta$ and $h_{32}(\theta) = \sin\theta$. Evaluating the LHS and RHS of equation (129), we get:

$$\begin{aligned}\text{LHS} &= \frac{1}{\sin 21}\frac{1.5\sin 22 - 1.1\sin 20}{2\pi/180} = 14.8439\\ \text{RHS} &= \frac{1.5 - 1.1}{2\pi/180} + \frac{1.5 + 1.1}{2}\cot 21 = 14.8458 \neq \text{LHS}.\end{aligned}$$

So how do we achieve a traceless shear tensor to machine round-off error?

Start by noting the following algebraic identities:

$$\left.\begin{aligned}\frac{v_1}{h_2}\partial_1 h_2 &= \frac{1}{h_2}\partial_1(h_2 v_1) - \partial_1 v_1;\\ \frac{v_1}{h_{31}}\partial_1 h_{31} &= \frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} v_1) - \frac{1}{h_2}\partial_1(h_2 v_1);\\ \frac{v_2}{h_{32}}\partial_2 h_{32} &= \frac{1}{h_{32}}\partial_2(h_{32} v_2) - \partial_2 v_2.\end{aligned}\right\} \tag{130}$$

Substitute equations (130) into equations (124) as appropriate, to get:

$$\left.\begin{aligned}
(\nabla\vec{v})_{11} &= \partial_1 v_1; \\
(\nabla\vec{v})_{22} &= \frac{1}{h_2}\partial_2 v_2 + \frac{1}{h_2}\partial_1(h_2 v_1) - \partial_1 v_1; \\
(\nabla\vec{v})_{33} &= \frac{1}{h_3}\partial_3 v_3 + \frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} v_1) - \frac{1}{h_2}\partial_1(h_2 v_1) + \frac{1}{h_2 h_{32}}\partial_2(h_{32} v_2) - \frac{1}{h_2}\partial_2 v_2,
\end{aligned}\right\} \quad (131)$$

giving us algebraically identical, but numerically only similar expressions to equations (124). However, this time using equations (131), we get:

$$\begin{aligned}
\sum_{i=1}^{3}(\nabla\vec{v})_{ii} &= \cancel{\partial_1 v_1} + \cancel{\frac{1}{h_2}\partial_2} v_2 + \cancel{\frac{1}{h_2}\partial_1(h_2 v_1)} - \cancel{\partial_1 v_1} + \frac{1}{h_3}\partial_3 v_3 + \frac{1}{h_2 h_{31}}\partial_1(h_2 h_{31} v_1) \\
&\quad - \cancel{\frac{1}{h_2}\partial_1(h_2 v_1)} + \frac{1}{h_2 h_{32}}\partial_2(h_{32} v_2) - \cancel{\frac{1}{h_2}\partial_2} v_2 = \nabla\cdot\vec{v},
\end{aligned}$$

to machine round off error, without having to rely on expressions (128) and (129). Note further that given face-centred velocities (as they are in a *staggered mesh* scheme like *ZEUS*), every term in equations (131) is naturally zone-centred without the need for two-point averages that are required if equations (124) are used. Two-point averages can be diffusive and, as we've seen, can lead to truncation errors much larger than the machine round-off limit.

This is a common strategy taken in numerics. Equations (131), which give slightly different estimates of $(\nabla\vec{v})_{ii}$ than equations (124), are used so that identities such as $\sum_i(\nabla\vec{v})_{ii} = \nabla\cdot\vec{v}$ are honoured to machine roundoff error. After all, who is to say whether the estimates of $(\nabla\vec{v})_{ii}$ afforded by equations (124) are any better or worse than those of equations (131)? Both are differenced estimates of differential quantities and both converge at the same rate to the differential quantity as δr and $\delta\theta \to 0$. The fact that equations (131) give a numerically traceless S whereas equations (124) do not is the discriminating factor, and makes equations (131) the more desirable of the two to be used in equation (126).

In terms of the numerical grid discussed in the *aside*, S_{ii} are zone-centred quantities. Meanwhile, the off-diagonal components, given by:

$$\begin{aligned}
S_{12} = S_{21} = (\nabla\vec{v})_{12} + (\nabla\vec{v})_{21} &= \partial_1 v_2 - \frac{v_2}{h_2}\partial_1 h_2 + \frac{1}{h_2}\partial_2 v_1 \\
&= h_2\,\partial_1\left(\frac{v_2}{h_2}\right) + \frac{1}{h_2}\partial_2 v_1; \qquad (132)
\end{aligned}$$

$$S_{13} = S_{31} = (\nabla\vec{v})_{13} + (\nabla\vec{v})_{31} = h_{31}\,\partial_1\left(\frac{v_3}{h_{31}}\right) + \frac{1}{h_3}\partial_3 v_1; \qquad (133)$$

$$S_{23} = S_{32} = (\nabla\vec{v})_{23} + (\nabla\vec{v})_{32} = \frac{h_{32}}{h_2}\partial_2\left(\frac{v_3}{h_{32}}\right) + \frac{1}{h_3}\partial_3 v_2, \qquad (134)$$

are naturally *edge-centred*, with S_{ij} located on the k-edge ($i \neq j \neq k$).

We now write down covariant expressions for the physical components of the vector constructs in equations (121) involving T. First, evaluate $(\nabla\cdot\mathsf{T})_1$ using equation (101) and the fact that T is symmetric (*e.g.*, $T_{21}=T_{12}$). Thus,

$$\begin{aligned}(\nabla\cdot\mathsf{T})_1 \;=\; &\frac{1}{h_2h_3}\left[\partial_1\!\left(\frac{h_2h_3}{h_1}T_{11}\right)+\partial_2(h_3T_{12})+\partial_3(h_2T_{13})\right]+\frac{1}{h_1^2}\big(\cancel{2}T_{11}\partial_1h_1-\cancel{T_{11}\partial_1h_1}\big)\\ &+\frac{1}{h_1h_2}\big(2T_{12}\partial_2h_1-T_{22}\partial_1h_2\big)+\frac{1}{h_1h_3}\big(2T_{13}\partial_3h_1-T_{33}\partial_1h_3\big)\end{aligned} \tag{135}$$

Now, a little algebra will show that:

$$\begin{aligned}\frac{1}{h_2h_3}\partial_1\!\left(\frac{h_2h_3}{h_1}T_{11}\right)+\frac{T_{11}}{h_1^2}\partial_1h_1 \;&=\; \frac{1}{h_1h_2h_3}\partial_1(h_2h_3T_{11});\\ \frac{1}{h_2h_3}\partial_2(h_3T_{12})+\frac{2T_{12}}{h_1h_2}\partial_2h_1 \;&=\; \frac{1}{h_1^2h_2h_3}\partial_2(h_1^2h_3T_{12});\quad\text{and}\\ \frac{1}{h_2h_3}\partial_3(h_2T_{13})+\frac{2T_{13}}{h_1h_3}\partial_3h_1 \;&=\; \frac{1}{h_1^2h_2h_3}\partial_3(h_1^2h_2T_{13}),\end{aligned}$$

and equation (135) simplifies to:

$$\begin{aligned}(\nabla\cdot\mathsf{T})_1 \;=\; \frac{1}{h_1h_2h_3}\bigg[&\partial_1(h_2h_3T_{11})+\frac{1}{h_1}\partial_2(h_1^2h_3T_{12})+\frac{1}{h_1}\partial_3(h_1^2h_2T_{13})\\ &-h_3T_{22}\partial_1h_2-h_2T_{33}\partial_1h_3\bigg].\end{aligned} \tag{136}$$

Next, since T is traceless ($T_{11}=-T_{22}-T_{33}$), we can write:

$$\begin{aligned}&\partial_1(h_2h_3T_{11})-h_3T_{22}\partial_1h_2-h_2T_{33}\partial_1h_3\\ &\qquad=\;-\partial_1(h_2h_3T_{22})-h_3T_{22}\partial_1h_2-\partial_1(h_2h_3T_{33})-h_2T_{33}\partial_1h_3\\ &\qquad=\;-h_2\partial_1(h_3T_{22})-2h_3T_{22}\partial_1h_2-h_3\partial_1(h_2T_{33})-2h_2T_{33}\partial_1h_3\\ &\qquad=\;-\frac{1}{h_2}\partial_1(h_2^2h_3T_{22})-\frac{1}{h_3}\partial_1(h_2h_3^2T_{33}),\end{aligned}$$

and equation (136) further reduces to:

$$(\nabla\cdot\mathsf{T})_1=\frac{1}{h_1h_2h_3}\left[-\frac{1}{h_2}\partial_1(h_2^2h_3T_{22})-\frac{1}{h_3}\partial_1(h_2h_3^2T_{33})+\frac{1}{h_1}\partial_2(h_1^2h_3T_{12})+\frac{1}{h_1}\partial_3(h_1^2h_2T_{13})\right]. \tag{137}$$

Permuting the indices, we get for the 2- and 3-components:

$$(\nabla\cdot\mathsf{T})_2=\frac{1}{h_1h_2h_3}\left[-\frac{1}{h_3}\partial_2(h_3^2h_1T_{33})-\frac{1}{h_1}\partial_2(h_3h_1^2T_{11})+\frac{1}{h_2}\partial_3(h_2^2h_1T_{23})+\frac{1}{h_2}\partial_1(h_2^2h_3T_{12})\right]; \tag{138}$$

$$(\nabla\cdot\mathsf{T})_3=\frac{1}{h_1h_2h_3}\left[-\frac{1}{h_1}\partial_3(h_1^2h_2T_{11})-\frac{1}{h_2}\partial_3(h_1h_2^2T_{22})+\frac{1}{h_3}\partial_1(h_3^2h_2T_{13})+\frac{1}{h_3}\partial_2(h_3^2h_1T_{23})\right]. \tag{139}$$

For Cartesian, cylindrical, or spherical polar coordinates, $h_1 = 1$, $h_2 = h_2(x_1)$, and $h_3 = h_{31}(x_1)h_{32}(x_2)$, so that equations (137), (138), and (139) become:

$$(\nabla\cdot\mathsf{T})_1 = -\frac{1}{h_2^2 h_{31}}\partial_1(h_2^2 h_{31} T_{22}) - \frac{1}{h_2 h_{31}^2}\partial_1(h_2 h_{31}^2 T_{33}) + \frac{1}{h_2 h_{32}}\partial_2(h_{32}T_{12}) + \frac{1}{h_3}\partial_3 T_{13}; \quad (140)$$

$$(\nabla\cdot\mathsf{T})_2 = -\frac{1}{h_2 h_{32}^2}\partial_2(h_{32}^2 T_{33}) - \frac{1}{h_2 h_{32}}\partial_2(h_{32}T_{11}) + \frac{1}{h_3}\partial_3 T_{23} + \frac{1}{h_2 h_{31}}\partial_1(h_2^2 h_{31} T_{12}); \quad (141)$$

$$(\nabla\cdot\mathsf{T})_3 = \underbrace{-\frac{1}{h_3}\partial_3(T_{11}) - \frac{1}{h_3}\partial_3(T_{22})}_{h_3^{-1}\partial_3 T_{33}} + \frac{1}{h_2 h_{31}^2}\partial_1(h_2 h_{31}^2 T_{13}) + \frac{1}{h_2 h_{32}^2}\partial_2(h_{32}^2 T_{23}). \quad (142)$$

Thinking in terms of a staggered numerical grid once again, since $T_{ij} \propto S_{ij}$, T_{ii} are zone-centred and T_{ij} are k-edge centred ($i \neq j \neq k$). Thus, with a little examination, one can see that every term in $(\nabla\cdot\mathsf{T})_i$ is naturally centred at the i-face without the need for two-point averaging, exactly where they are needed to accelerate the i-face centred v_i. This is the principle strength of a properly staggered mesh: vector components often naturally land where the equations need them to be, without the need of averaging.

Next, from equation (41), we have:

$$\begin{aligned}
\mathsf{T}:\nabla\vec{v} &= \sum_{ij} T_{ij}(\nabla\vec{v})_{ij} \\
&= T_{11}\nabla_1 v_1 + T_{22}\nabla_2 v_2 + T_{33}\nabla_3 v_3 \\
&\quad + T_{12}(\nabla_1 v_2 + \nabla_2 v_1) + T_{13}(\nabla_1 v_3 + \nabla_3 v_1) + T_{23}(\nabla_2 v_3 + \nabla_3 v_2) \\
&= T_{11}\tfrac{1}{2}S_{11} + T_{22}\tfrac{1}{2}S_{22} + T_{33}\tfrac{1}{2}S_{33} + \tfrac{1}{3}\big(\underbrace{T_{11}+T_{22}+T_{33}}_{=\,\mathrm{Tr}(\mathsf{T})\,=\,0}\big)\nabla\cdot\vec{v} \\
&\quad + T_{12}S_{12} + T_{13}S_{13} + T_{23}S_{23},
\end{aligned}$$

exploiting the symmetry of T and using equations (126), (132), (133) and (134). Thus,

$$\mathsf{T}:\nabla\vec{v} = \frac{1}{2}\sum_{ij} T_{ij}S_{ij} = \frac{\mu^*}{2}\sum_{ij} S_{ij}^2, \quad (143)$$

where $\mu^* \equiv \mu + \mu_{\rm l} + \mu_{\rm q}$. Alternately, we can write:

$$\begin{aligned}
\mathsf{T}:\nabla\vec{v} &= \sum_{ij} T_{ij}(\nabla\vec{v})_{ij} = \mu^*\sum_i S_{ii}\nabla_i v_i + \mu^*\sum_{i\neq j} S_{ij}\nabla_i v_j \\
&= \mu^*\sum_i \big(2\nabla_i v_i - \tfrac{2}{3}\nabla\cdot\vec{v}\big)\nabla_i v_i + \mu^*\sum_{i\neq j}\big(\nabla_i v_j + \nabla_j v_i\big)\nabla_i v_j \\
&= 2\mu^*\sum_i(\nabla_i v_i)^2 - \frac{2\mu^*}{3}\underbrace{\nabla\cdot\vec{v}}_{\sum_j\nabla_j v_j}\sum_i\nabla_i v_i + \frac{\mu^*}{2}\sum_{i\neq j}\big(\nabla_i v_j + \nabla_j v_i\big)^2 \\
&= \frac{2\mu^*}{3}\Big(3(\nabla_1 v_1)^2 + 3(\nabla_2 v_2)^2 + 3(\nabla_3 v_3)^2 - (\nabla_1 v_1)^2 - (\nabla_2 v_2)^2 - (\nabla_3 v_3)^2
\end{aligned}$$

$$
\begin{aligned}
&\qquad - 2\nabla_1 v_1 \nabla_2 v_2 - 2\nabla_2 v_2 \nabla_3 v_3 - 2\nabla_3 v_3 \nabla_1 v_1\Big) + \frac{\mu^*}{2}\sum_{i\neq j}\big(\nabla_i v_j + \nabla_j v_i\big)^2 \\
&= \frac{2\mu^*}{3}\Big((\nabla_1 v_1)^2 - 2\nabla_1 v_1 \nabla_2 v_2 + (\nabla_2 v_2)^2 + (\nabla_2 v_2)^2 - 2\nabla_2 v_2 \nabla_3 v_3 + (\nabla_3 v_3)^2 \\
&\qquad + (\nabla_3 v_3)^2 - 2\nabla_3 v_3 \nabla_1 v_1 + (\nabla_1 v_1)^2\Big) + \frac{\mu^*}{2}\sum_{i\neq j}\big(\nabla_i v_j + \nabla_j v_i\big)^2
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad \mathsf{T} : \nabla\vec{v} \;=\; & \frac{2\mu^*}{3}\Big(\big(\nabla_1 v_1 - \nabla_2 v_2\big)^2 + \big(\nabla_2 v_2 - \nabla_3 v_3\big)^2 + \big(\nabla_3 v_3 - \nabla_1 v_1\big)^2\Big) \\
& + \mu^*\Big(\big(\nabla_1 v_2 + \nabla_2 v_1\big)^2 + \big(\nabla_2 v_3 + \nabla_3 v_2\big)^2 + \big(\nabla_3 v_1 + \nabla_1 v_3\big)^2\Big)
\end{aligned}
\tag{144}
$$

Stone and Norman (1992, ApJS, 80, p. 789) state that TW claim that in the context of a numerical algorithm, *representing the viscous energy generation term as a sum of squares as in equation (144) is an absolute necessity, because all other formulations eventually lead to instabilities*. Since equations (143) and (144) differ only in the order in which terms are added, they must yield results identical to machine round-off error, and I cannot understand why such a claim would be made. Indeed, equation (143) should have the same stability properties as equation (144).

Finally, from equation (114), we have:

$$
\begin{aligned}
\nabla\cdot(\mathsf{T}\cdot\vec{v}) \;=\; & \frac{1}{h_2 h_{31}}\partial_1\Big(h_2 h_{31}\big(T_{11}v_1 + T_{12}v_2 + T_{13}v_3\big)\Big) \\
& + \frac{1}{h_2 h_{32}}\partial_2\Big(h_{32}\big(T_{12}v_1 + T_{22}v_2 + T_{23}v_3\big)\Big) + \frac{1}{h_3}\partial_3\big(T_{13}v_1 + T_{23}v_2 + T_{33}v_3\big).
\end{aligned}
\tag{145}
$$

Centring equation (145) is awkward since T_{ij} is zone centred for $i = j$ and edge-centred for $i \neq j$, while v_j are face-centred. This is an example where the staggered mesh does not seem to help, and numerous and seemingly unnatural 2- and 4-point averages are necessary to evaluate all products in $\nabla\cdot(\mathsf{T}\cdot\vec{v})$ using equation (145). Still, it is this form that the term is a perfect divergence and thus most useful in representing the total energy equation in a conservative fashion.

Alternately, using the identity:

$$
\nabla\cdot(\mathsf{T}\cdot\vec{v}) \;=\; \mathsf{T} : \nabla\vec{v} + (\nabla\cdot\mathsf{T})\cdot\vec{v}, \tag{146}
$$

(Theorem 5.2, §5.9) factors in each product are now co-spatial, though 2- and 4-point averages are still needed to bring some—not all—of the terms to the zone centres. While the centring of the RHS of equation (146) may seem more natural than the LHS, the RHS is not in conservative form, and this deficit may trump the centring advantage. This can be determined only by direct experimentation.

Last point. Note that in Cartesian coordinates, ignoring all diagonal components, dropping the $\nabla\cdot\vec{v}$ terms in equation (126), and letting $\nabla\cdot\vec{v}$ in equation (123) $\rightarrow \partial_i v_i$ in the i direction, equations (137)–(139) reduce to:

$$
\nabla\cdot\mathsf{T} \;=\; 2\big(\partial_x(\mu^*\partial_x v_x),\, \partial_y(\mu^*\partial_y v_y),\, \partial_z(\mu^*\partial_z v_z)\big), \tag{147}
$$

equation (145) reduces to:

$$\nabla \cdot (\mathsf{T} \cdot \vec{v}) \;=\; 2\big(\partial_x(v_x\mu^*\partial_x v_x) + \partial_y(v_y\mu^*\partial_y v_y) + \partial_z(v_z\mu^*\partial_z v_z)\big), \tag{148}$$

while equation (143) reduces to:

$$\mathsf{T} : \nabla\vec{v} \;=\; 2\mu^*\big((\partial_x v_x)^2 + (\partial_y v_y)^2 + (\partial_z v_z)^2\big). \tag{149}$$

Equations (147)–(149) are the vNR expressions for the artificial viscosity in the subroutine `viscous` in *ZEUS*, when $\mu = 0$ and μ^* is taken to be (equation 123):

$$\mu_{\rm l} + \mu_{\rm q} \;=\; \underbrace{2q_1}_{\texttt{qlin}} l\rho c_{\rm s} - \underbrace{2q_2}_{\texttt{qcon}} l^2\rho \min\big(0, \partial_i v_i\big).$$

$\sim$